

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
10 May 2001 (10.05.2001)

PCT

(10) International Publication Number  
WO 01/32888 A2(51) International Patent Classification<sup>7</sup>: C12N 15/54,  
9/10, 9/12, A01K 67/00, C07K 16/40, A61K 38/45, G01N  
33/50, C12Q 1/68

(21) International Application Number: PCT/US00/30485

(22) International Filing Date:  
2 November 2000 (02.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/163,595 4 November 1999 (04.11.1999) US(63) Related by continuation (CON) or continuation-in-part  
(CIP) to earlier application:  
US 60/163,595 (CIP)  
Filed on 4 November 1999 (04.11.1999)(71) Applicant (for all designated States except US): INCYTE  
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo  
Alto, CA 94030 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom  
[CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).  
YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale,  
CA 94087 (US). HILLMAN, Jennifer, L. [US/US]; 230  
Monroe Drive, #17, Mountain View, CA 94040 (US).  
LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara,  
CA 95054 (US). BANDMAN, Olga [US/US]; 366 Anna  
Avenue, Mountain View, CA 94043 (US). PATTERSON,Chandra [US/US]; 490 Sherwood Way, #1, Menlo Park,  
CA 94025 (US). SHIH, Leo, L. [US/US]; Apartment  
B, 1081 Tanland Drive, Palo Alto, CA 94303 (US). AZ-  
IMZAI, Yalda [US/US]; 5518 Boulder Canyon Drive,  
Castro Valley, CA 94552 (US). LU, Dyung, Aina, M.  
[US/US]; 233 Coy Drive, San Jose, CA 95123 (US).  
BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road,  
San Leandro, CA 94577 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,  
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GI, GM,  
KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).**Published:***Without international search report and to be republished  
upon receipt of that report.**For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

WO 01/32888 A2

(54) Title: HUMAN TRANSFERASE MOLECULES

(57) Abstract: The invention provides human transferase molecules (HTFS) and polynucleotides which identify and encode HTFS. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HTFS.

## HUMAN TRANSFERASE MOLECULES

### TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human transferase molecules  
5 and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative disorders and immune system disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of human transferase molecules.

### BACKGROUND OF THE INVENTION

10       Transferases are enzymes that catalyze the transfer of molecular groups. The reaction may involve an oxidation, reduction, or cleavage of covalent bonds, and is often specific to a substrate or to particular sites on a type of substrate. Transferases participate in reactions essential to such functions as synthesis and degradation of cell components, regulation of cell functions including cell signaling, cell proliferation, inflammation, apoptosis, secretion and excretion. Transferases are  
15 involved in key steps in disease processes involving these functions. Transferases are frequently classified according to the type of group transferred. For example, methyl transferases transfer one-carbon methyl groups, amino transferases transfer nitrogenous amino groups, and similarly denominated enzymes transfer aldehyde or ketone, acyl, glycosyl, alkyl or aryl, isoprenyl, saccharyl, phosphorous-containing, sulfur-containing, or selenium-containing groups, as well as small enzymatic  
20 groups such as Coenzyme A.

Acyl transferases include peroxisomal carnitine octanoyl transferase, which is involved in the fatty acid beta-oxidation pathway, and mitochondrial carnitine palmitoyl transferases, involved in fatty acid metabolism and transport. Choline O-acetyl transferase catalyzes the biosynthesis of the neurotransmitter acetylcholine.

25       Amino transferases play key roles in protein synthesis and degradation, and they contribute to other processes as well. For example, the amino transferase 5-aminolevulinic acid synthase catalyzes the addition of succinyl-CoA to glycine, the first step in heme biosynthesis. Other amino transferases participate in pathways important for neurological function and metabolism. For example, glutamine-phenylpyruvate amino transferase, also known as glutamine transaminase K (GTK), catalyzes several  
30 reactions with a pyridoxal phosphate cofactor. GTK catalyzes the reversible conversion of L-glutamine and phenylpyruvate to 2-oxoglutaramate and L-phenylalanine. Other amino acid substrates for GTK include L-methionine, L-histidine, and L-tyrosine. GTK also catalyzes the conversion of kynurenine to kynurenic acid, a tryptophan metabolite that is an antagonist of the N-methyl-D-aspartate (NMDA) receptor in the brain and may exert a neuromodulatory function. Alteration of the kynurenine metabolic  
35 pathway may be associated with several neurological disorders. GTK also plays a role in the

metabolism of halogenated xenobiotics conjugated to glutathione, leading to nephrotoxicity in rats and neurotoxicity in humans. GTK is expressed in kidney, liver, and brain. Both human and rat GTKs contain a putative pyridoxal phosphate binding site. (ExpASY ENZYME: EC 2.6.1.64; Perry, S.J. et al. (1993) Mol. Pharmacol. 43:660-665; Perry, S. et al. (1995) FEBS Lett. 360:277-280; and Alberati-Giani, D. et al. (1995) J. Neurochem. 64:1448-1455.) A second amino transferase associated with this pathway is kynurenine/ $\alpha$ -aminoadipate amino transferase (AadAT). AadAT catalyzes the reversible conversion of  $\alpha$ -aminoadipate and  $\alpha$ -ketoglutarate to  $\alpha$ -ketoadipate and L-glutamate during lysine metabolism. AadAT also catalyzes the transamination of kynurenine to kynurenic acid. A cytosolic AadAT is expressed in rat kidney, liver, and brain. (Nakatani, Y. et al. (1970) Biochim. Biophys. Acta 198:219-228; Buchli, R. et al. (1995) J. Biol. Chem. 270:29330-29335).

Glycosyl transferases include the mammalian UDP-glucouronosyl transferases, a family of membrane-bound microsomal enzymes catalyzing the transfer of glucouronic acid to lipophilic substrates in reactions that play important roles in detoxification and excretion of drugs, carcinogens, and other foreign substances. Another mammalian glycosyl transferase, mammalian UDP-galactose-ceramide galactosyl transferase, catalyzes the transfer of galactose to ceramide in the synthesis of galactocerebrosides in myelin membranes of the nervous system. The UDP-glycosyl transferases share a conserved signature domain of about 50 amino acid residues (PROSITE: PDOC00359, <http://expasy.hcuge.ch/sprot/prosite.html>).

Methyl transferases are involved in a variety of pharmacologically important processes.

Nicotinamide N-methyl transferase catalyzes the N-methylation of nicotinamides and other pyridines, an important step in the cellular handling of drugs and other foreign compounds. Phenylethanolamine N-methyl transferase catalyzes the conversion of noradrenalin to adrenalin. 6-O-methylguanine-DNA methyl transferase reverses DNA methylation, an important step in carcinogenesis. Uroporphyrin-III C-methyl transferase, which catalyzes the transfer of two methyl groups from S-adenosyl-L-methionine to uroporphyrinogen III, is the first-specific-enzyme in the biosynthesis of cobalamin, a dietary enzyme whose uptake is deficient in pernicious anemia. Protein-arginine methyl transferases catalyze the posttranslational methylation of arginine residues in proteins, resulting in the mono- and dimethylation of arginine on the guanidino group. Substrates include histones, myelin basic protein, and heterogeneous nuclear ribonucleoproteins involved in mRNA processing, splicing, and transport.

Protein-arginine methyl transferase interacts with proteins upregulated by mitogens, with proteins involved in chronic lymphocytic leukemia, and with interferon, suggesting an important role for methylation in cytokine receptor signaling (Lin, W.-J. et al. (1996) J. Biol. Chem. 271:15034-15044; Abramovich, C. et al. (1997) EMBO J. 16:260-266; and Scott, H.S. et al. (1998) Genomics 48:330-340.)

Phospho-transferases catalyze the transfer of high-energy phosphate groups and are important in energy-requiring and -releasing reactions. The metabolic enzyme creatine kinase catalyzes the reversible phosphate transfer between creatine/creatine phosphate and ATP/ADP. Glycocyamine kinase catalyzes phosphate transfer from ATP to guanidoacetate, and arginine kinase catalyzes phosphate transfer from ATP to arginine. A cysteine-containing active site is conserved in this family (PROSITE: PDOC00103).

Prenyl transferases are heterodimers, consisting of an alpha and a beta subunit, that catalyze the transfer of an isoprenyl group. An example of a prenyl transferase is the mammalian protein farnesyl transferase. The alpha subunit of farnesyl transferase consists of 5 repeats of 34 amino acids each, with each repeat containing an invariant tryptophan (PROSITE: PDOC00703).

Saccharyl transferases are glycosylating enzymes involved in a variety of metabolic processes. Oligosaccharyl transferase-48, for example, is a receptor for advanced glycation endproducts. Accumulation of these endproducts is observed in vascular complications of diabetes, macrovascular disease, renal insufficiency, and Alzheimer's disease (Thornalley, P.J. (1998) Cell Mol. Biol. (Noisy-Le-Grand) 44:1013-1023).

Coenzyme A (CoA) transferase catalyzes the transfer of CoA between two carboxylic acids. Succinyl CoA:3-oxoacid CoA transferase, for example, transfers CoA from succinyl-CoA to a recipient such as acetoacetate. Acetoacetate is essential to the metabolism of ketone bodies, which accumulate in tissues affected by metabolic disorders such as diabetes (PROSITE: PDOC00980).

The discovery of new human transferase molecules and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative disorders and immune system disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of human transferase molecules.

## SUMMARY OF THE INVENTION

The invention features purified polypeptides, human transferase molecules, referred to collectively as "HTFS" and individually as "HTFS-1," "HTFS-2," "HTFS-3," "HTFS-4," "HTFS-5," "HTFS-6," "HTFS-7," "HTFS-8," "HTFS-9," "HTFS-10," "HTFS-11," "HTFS-12," "HTFS-13," "HTFS-14," "HTFS-15," "HTFS-16," "HTFS-17," "HTFS-18," "HTFS-19," "HTFS-20," "HTFS-21," "HTFS-22," "HTFS-23," "HTFS-24," "HTFS-25," "HTFS-26," "HTFS-27," "HTFS-28," "HTFS-29," "HTFS-30," "HTFS-31," "HTFS-32," "HTFS-33," "HTFS-34," "HTFS-35," "HTFS-36," "HTFS-37," "HTFS-38," "HTFS-39," "HTFS-40," "HTFS-41," and "HTFS-42." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the



group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an

- 5 immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-42.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least  
10 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. In one alternative, the polynucleotide encodes a polypeptide selected  
15 from the group consisting of SEQ ID NO:1-42. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:43-84.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group  
20 consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. In one alternative, the invention provides a cell transformed with the  
25 recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90%  
30 sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide

comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid  
5 sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

10 The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e)  
15 an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of  
20 SEQ ID NO:43-84, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and c) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide  
25 in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

30 The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, c) a

polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional HTFS, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional HTFS, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment

of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In  
5 another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional HTFS, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino  
10 acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic fragment of an amino acid  
15 sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a)  
20 an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and d) an immunogenic  
25 fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the  
test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a  
30 compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:43-84, the method comprising a)  
35 exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

#### BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HTFS.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of HTFS.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding HTFS were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these  
5 may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a  
10 reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although  
15 any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is  
20 not entitled to antedate such disclosure by virtue of prior invention.

## DEFINITIONS

"HTFS" refers to the amino acid sequences of substantially purified HTFS obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

25 The term "agonist" refers to a molecule which intensifies or mimics the biological activity of HTFS. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HTFS either by directly interacting with HTFS or by acting on components of the biological pathway in which HTFS participates.

An "allelic variant" is an alternative form of the gene encoding HTFS. Allelic variants may  
30 result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in

a given sequence.

"Altered" nucleic acid sequences encoding HTFS include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HTFS or a polypeptide with at least one functional characteristic of HTFS. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HTFS, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HTFS. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HTFS. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HTFS is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of HTFS. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HTFS either by directly interacting with HTFS or by acting on components of the biological pathway in which HTFS participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind HTFS polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or

synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic HTFS, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HTFS or fragments of HTFS may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be



deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
20	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
25	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
30	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of HTFS or the polynucleotide encoding HTFS which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:43-84 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:43-84, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:43-84 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:43-84 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:43-84 and the region of SEQ ID NO:43-84 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-42 is encoded by a fragment of SEQ ID NO:43-84. A fragment of SEQ ID NO:1-42 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-42. For example, a fragment of SEQ ID NO:1-42 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-42. The precise length of a fragment of SEQ ID NO:1-42 and the region of SEQ ID NO:1-42 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon

(e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

5       The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

10       Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191.

15       For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

20       Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 11*

5 *Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such  
10 lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in  
15 a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment  
20 methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment  
25 program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

30 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Open Gap: 11 and Extension Gap: 1 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 3*

*Filter: on*

5       Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length  
10 supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

      "Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

15       The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

      "Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific  
20 hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for  
25 annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

30       Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions

for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C<sub>0</sub>t or R<sub>0</sub>t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of HTFS which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of HTFS which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of HTFS. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HTFS.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an HTFS may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of HTFS.

"Probe" refers to nucleic acid sequences encoding HTFS, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques



such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

5 Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs).

10 Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and  
15 other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

20 The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HTFS, or fragments thereof, or HTFS itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a  
25 protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will  
30 reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook, J. et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater

sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may

5 possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in

10 which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999)

15 set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

#### THE INVENTION

The invention is based on the discovery of new human transferase molecules (HTFS), the

20 polynucleotides encoding HTFS, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative disorders and immune system disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HTFS. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which

25 nucleic acids encoding each HTFS were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In some cases, GenBank sequence identifiers are also shown in column 5. The Incyte clones and GenBank cDNA sequences, where indicated, in column 5 were used to assemble the consensus nucleotide

30 sequence of each HTFS and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6

35 shows homologous sequences as identified by BLAST analysis along with relevant citations, all of

which are expressly incorporated by reference herein in their entirety; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

5           The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HTFS. The first column of Table 3 lists the nucleotide SEQ ID NOs. Fragments of these nucleotide sequences are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:43-84 and to distinguish between SEQ ID NO:43-84 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for  
10       example, as immunogenic peptides. Column 2 lists tissue categories which express HTFS as a fraction of total tissues expressing HTFS. Column 3 lists diseases, disorders, or conditions associated with those tissues expressing HTFS as a fraction of total tissues expressing HTFS. Column 4 lists the vectors used to subclone each cDNA library.

          The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries  
15       from which cDNA clones encoding HTFS were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

          SEQ ID NO:44 maps to chromosome 1 within the interval from 170.1 to 186.4 centiMorgans. SEQ ID NO:46 maps to chromosome 11 within the interval from 58.2 to 59.5 centiMorgans. SEQ ID  
20       NO:48 maps to chromosome 11 within the interval from 67.4 to 70.9 centiMorgans. SEQ ID NO:49 maps to chromosome 21 within the interval from 51.6 centiMorgans to the q-terminus. SEQ ID NO:52 maps to chromosome 3 within the interval from 63.3 to 77.4 centiMorgans. SEQ ID NO:59 maps to chromosome 20 within the interval from 50.2 to 53.6 centiMorgans and to chromosome 12 within the interval from 113.3 to 118.9 centiMorgans. SEQ ID NO:60 maps to chromosome 12  
25       within the interval from 62.7 to 70.6 centiMorgans. SEQ ID NO:62 maps to chromosome 11 within the interval from 62.5 to 70.9 centiMorgans. SEQ ID NO:68 maps to chromosome 11 within the interval from 70.9 to 72.1 centiMorgans. SEQ ID NO:78 maps to chromosome 23 within the interval from 94.4 to 97.4 centiMorgans and to chromosome 2 within the interval from 272.5 centiMorgans to the q-terminus. SEQ ID NO:85 maps to chromosome 5 within the interval from 5.5 to 21.5  
30       centiMorgans, to chromosome 17 within the interval from 53.9 to 62.9 centiMorgans, and to chromosome 12 within the interval from 84.7 to 92.5 centiMorgans. SEQ ID NO:86 maps to chromosome 6 within the interval from 42.0 to 45.4 centiMorgans, to chromosome 11 within the interval from 58.2 to 59.5 centiMorgans, and to chromosome 16 within the interval from 88.1 to 92.6 centiMorgans.

35           The invention also encompasses HTFS variants. A preferred HTFS variant is one which has at

least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the HTFS amino acid sequence, and which contains at least one functional or structural characteristic of HTFS.

5 The invention also encompasses polynucleotides which encode HTFS. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:43-84, which encodes HTFS. The polynucleotide sequences of SEQ ID NO:43-84, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

10 The invention also encompasses a variant of a polynucleotide sequence encoding HTFS. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HTFS. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:43-84 which has at  
15 least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:43-84. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HTFS.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic  
20 code, a multitude of polynucleotide sequences encoding HTFS, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally  
25 occurring HTFS, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HTFS and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HTFS under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HTFS or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring  
30 codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HTFS and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced

from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HTFS and HTFS derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HTFS or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:43-84 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (PE Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HTFS may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids*

Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown  
5 sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using  
10 commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been  
15 size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the  
20 size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems), and the entire process  
25 from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HTFS may be cloned in recombinant DNA molecules that direct expression of HTFS, or  
30 fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HTFS.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HTFS-encoding sequences for a variety of purposes including, but not

limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of HTFS, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding HTFS may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, HTFS itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of HTFS, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing.



(See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active HTFS, the nucleotide sequences encoding HTFS or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HTFS. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HTFS. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HTFS and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HTFS and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HTFS. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105; The

McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

10 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HTFS. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HTFS can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPT1 plasmid (Life Technologies). Ligation of sequences encoding HTFS into the vector's multiple cloning site  
15 disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HTFS are needed, e.g. for the production of antibodies,  
20 vectors which direct high level expression of HTFS may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HTFS. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such  
25 vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, supra; and Scorer, supra.)

Plant systems may also be used for expression of HTFS. Transcription of sequences encoding HTFS may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in  
30 combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, supra; Broglie, supra; and Winter, supra.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp.  
35 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HTFS may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain  
5 infective virus which expresses HTFS in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of  
10 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HTFS in cell lines is preferred. For example, sequences encoding HTFS can be transformed into cell  
15 lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the  
20 introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apv* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232;  
25 Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.)  
30 Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of

transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HTFS is inserted within a marker gene sequence, transformed cells containing sequences encoding HTFS can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HTFS under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HTFS and that express HTFS may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HTFS using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HTFS is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HTFS include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HTFS, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates,

cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HTFS may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HTFS may be designed to contain signal sequences which direct secretion of HTFS through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HTFS may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HTFS protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HTFS activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HTFS encoding sequence and the heterologous protein sequence, so that HTFS may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HTFS may be achieved in

vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

5 HTFS of the present invention or fragments thereof may be used to screen for compounds that specifically bind to HTFS. At least one and up to a plurality of test compounds may be screened for specific binding to HTFS. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of  
10 HTFS, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which HTFS binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for  
15 these compounds involves producing appropriate cells which express HTFS, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing HTFS or cell membrane fractions which contain HTFS are then contacted with a test compound and binding, stimulation, or inhibition of activity of either HTFS or the compound is analyzed.

20 An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with HTFS, either in solution or affixed to a solid support, and detecting the binding of HTFS to the compound.

Alternatively, the assay may detect or measure binding of a test compound in the presence of a  
25 labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

HTFS of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of HTFS. Such compounds may include agonists, antagonists, or partial or  
30 inverse agonists. In one embodiment, an assay is performed under conditions permissive for HTFS activity, wherein HTFS is combined with at least one test compound, and the activity of HTFS in the presence of a test compound is compared with the activity of HTFS in the absence of the test compound. A change in the activity of HTFS in the presence of the test compound is indicative of a compound that modulates the activity of HTFS. Alternatively, a test compound is combined with an  
35 in vitro or cell-free system comprising HTFS under conditions suitable for HTFS activity, and the

assay is performed. In either of these assays, a test compound which modulates the activity of HTFS may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding HTFS or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding HTFS may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) *Science* 282:1145-1147).

Polynucleotides encoding HTFS can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding HTFS is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress HTFS, e.g., by secreting HTFS in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

#### THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HTFS and human transferase molecules. In addition, the expression of HTFS is

closely associated with proliferating tissues and inflammation. Therefore, HTFS appears to play a role in cell proliferative disorders and immune system disorders. In the treatment of disorders associated with increased HTFS expression or activity, it is desirable to decrease the expression or activity of HTFS. In the treatment of disorders associated with decreased HTFS expression or activity, it is desirable to increase the expression or activity of HTFS.

Therefore, in one embodiment, HTFS or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTFS. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an immune system disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma.

In another embodiment, a vector capable of expressing HTFS or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTFS including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified HTFS in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a



disorder associated with decreased expression or activity of HTFS including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HTFS may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTFS including, but not limited to, those listed above.

In a further embodiment, an antagonist of HTFS may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HTFS. Examples of such disorders include, but are not limited to, those cell proliferative disorders and immune system disorders described above. In one aspect, an antibody which specifically binds HTFS may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express HTFS.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HTFS may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HTFS including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HTFS may be produced using methods which are generally known in the art. In particular, purified HTFS may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HTFS. Antibodies to HTFS may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HTFS or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG

(bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HTFS have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are  
5 identical to a portion of the amino acid sequence of the natural protein. Short stretches of HTFS amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HTFS may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited  
10 to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the  
15 splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HTFS-specific single  
20 chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in  
25 the literature. (See, e.g., Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

Antibody fragments which contain specific binding sites for HTFS may also be generated. For example, such fragments include, but are not limited to,  $F(ab')_2$  fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the  $F(ab')_2$   
30 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) *Science* 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either

polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HTFS and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HTFS epitopes is generally used, but a competitive binding assay may also be employed

5 (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HTFS. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of HTFS-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined

10 for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HTFS epitopes, represents the average affinity, or avidity, of the antibodies for HTFS. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular HTFS epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the HTFS-antibody complex must

15 withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HTFS, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

20 The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of HTFS-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for

25 antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding HTFS, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA,

30 PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding HTFS. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HTFS. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense

sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding HTFS may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in HTFS expression or regulation causes disease, the expression of HTFS from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in HTFS are treated by constructing mammalian expression vectors encoding HTFS and introducing these vectors by mechanical means into HTFS-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the

use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. R  capon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of HTFS include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). HTFS may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding HTFS from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to HTFS expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding HTFS under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and

A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference.

- 5 Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4<sup>+</sup> T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

- 10 In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding HTFS to cells which have one or more genetic abnormalities with respect to the expression of HTFS. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas  
15 (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544; and Verma, I.M. and N. Somia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

- 20 In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding HTFS to target cells which have one or more genetic abnormalities with respect to the expression of HTFS. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing HTFS to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with  
25 ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV  
30 d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus

sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

5 In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding HTFS to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid  
10 proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for HTFS into the alphavirus genome in place of the capsid-coding region results in the production of a large number of HTFS-coding RNAs and the synthesis of high levels of HTFS in vector transduced cells. While alphavirus  
15 infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of HTFS into a variety of cell types. The specific transduction of a subset of cells in a  
20 population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can  
25 be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A  
30 complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example,

engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HTFS.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HTFS. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding HTFS. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased HTFS expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding HTFS may be therapeutically useful, and in the treatment of disorders associated with decreased HTFS expression or activity, a compound which specifically promotes expression of the polynucleotide encoding HTFS may be therapeutically useful.



At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding HTFS is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding HTFS are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding HTFS. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient.

Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of HTFS, antibodies to HTFS, and mimetics, agonists, antagonists, or inhibitors of HTFS.

5       The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form.

10       These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle  
15       injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

20       Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising HTFS or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, HTFS or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to  
25       transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route  
30       of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HTFS or fragments thereof, antibodies of HTFS, and agonists, antagonists or inhibitors of HTFS, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by

- standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.
- 10 The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy.
- 15 Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu\text{g}$  to 100,000  $\mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

- 20 Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

## DIAGNOSTICS

- In another embodiment, antibodies which specifically bind HTFS may be used for the diagnosis of disorders characterized by expression of HTFS, or in assays to monitor patients being treated with HTFS or agonists, antagonists, or inhibitors of HTFS. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HTFS include methods which utilize the antibody and a label to detect HTFS in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.
- 25
- 30

A variety of protocols for measuring HTFS, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HTFS expression. Normal or standard values for HTFS expression are established by combining body fluids or cell extracts taken

from normal mammalian subjects, for example, human subjects, with antibody to HTFS under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HTFS expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation  
5 between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HTFS may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of HTFS may be correlated with  
10 disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HTFS, and to monitor regulation of HTFS levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HTFS or closely related molecules may be used to identify nucleic acid sequences which encode HTFS. The specificity of the probe, whether it is made  
15 from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding HTFS, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50%  
20 sequence identity to any of the HTFS encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:43-84 or from genomic sequences including promoters, enhancers, and introns of the HTFS gene.

Means for producing specific hybridization probes for DNAs encoding HTFS include the cloning of polynucleotide sequences encoding HTFS or HTFS derivatives into vectors for the  
25 production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

30 Polynucleotide sequences encoding HTFS may be used for the diagnosis of disorders associated with expression of HTFS. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including  
35 adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in

particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an immune system disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma. The polynucleotide sequences encoding HTFS may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HTFS expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HTFS may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HTFS may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HTFS in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HTFS, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HTFS, under conditions suitable for hybridization or amplification.

Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HTFS may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HTFS, or a fragment of a polynucleotide complementary to the polynucleotide encoding HTFS, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding HTFS may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding HTFS are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as

DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of HTFS include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplax, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for HTFS, or HTFS or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number

5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples



are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for HTFS to quantify the levels of HTFS expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoza, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid

degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding HTFS may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding HTFS on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HTFS, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HTFS and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HTFS, or fragments thereof, and washed. Bound HTFS is then detected by methods well known in the art. Purified HTFS can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HTFS specifically compete with a test compound for binding HTFS. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HTFS.

In additional embodiments, the nucleotide sequences which encode HTFS may be used in any

molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Serial No. 60/163,595, are hereby expressly incorporated by reference.

10

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid

(Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

5 Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid  
10 purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal  
15 cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

## III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.  
20 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI  
25 PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA  
30 sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions,

references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:43-84. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related

molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar.

The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HTFS occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

#### V. Chromosomal Mapping of HTFS Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:43-84 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:43-84 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for

Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

The genetic map locations of SEQ ID NO:44, 46, 48, 49, 52, 59, 60, 62, 68, 78, 85, and 86 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:59, 78, 85, and 86, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:59, 78, 85, and 86 were assembled into their respective clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

#### VI. Extension of HTFS Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:43-84 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2:



94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, the polynucleotide sequences of SEQ ID NO:43-84 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

## VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:43-84 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is

specifically described, essentially the same procedure is used with larger nucleotide fragments.

Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ - $^{32}$ P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing  $10^7$  counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

- 10 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

#### VIII. Microarrays

- The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schna (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schna, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

- Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser

desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

5 Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ $\mu$ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ $\mu$ l RNase inhibitor, 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dTTP, 40  $\mu$ M dCTP, 40  $\mu$ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)<sup>+</sup> RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14  $\mu$ l 5X SSC/0.2% SDS.

20 Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5  $\mu$ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average

concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water.

- 5 Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in 0.2% SDS and distilled water as before.

#### Hybridization

- Hybridization reactions contain 9  $\mu$ l of sample mixture consisting of 0.2  $\mu$ g each of Cy3 and  
10 Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  $\mu$ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for  
15 about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

#### Detection

- Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines  
20 at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

- 25 In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is  
30 typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

- The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that  
35 location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples

from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

- 5       The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and  
10   measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

- A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used  
15   for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

#### IX. Complementary Polynucleotides

- Sequences complementary to the HTFS-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HTFS. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with  
20   smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HTFS. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HTFS-encoding transcript.

#### 25   X. Expression of HTFS

- Expression and purification of HTFS is achieved using bacterial or virus-based expression systems. For expression of HTFS in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the  
30   T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HTFS upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HTFS in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as  
35   baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HTFS

by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HTFS is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HTFS at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified HTFS obtained by these methods can be used directly in the assays shown in Examples XI and XV.

#### **XI. Demonstration of HTFS Activity**

Galactosyltransferase activity is determined by measuring the transfer of galactose from UDP-galactose to a GlcNAc-terminated oligosaccharide chain in a radioactive assay. (Kolbinger, F. et al. (1998) J. Biol. Chem. 273:58-65.) The HTFS sample is incubated with 14  $\mu$ l of assay stock solution (180 mM sodium cacodylate, pH 6.5, 1 mg/ml bovine serum albumin, 0.26 mM UDP-galactose, 2  $\mu$ l of UDP-[ $^3$ H]galactose), 1  $\mu$ l of  $\text{MnCl}_2$  (500 mM), and 2.5  $\mu$ l of  $\text{GlcNAc}\beta\text{O}-(\text{CH}_2)_4-\text{CO}_2\text{Me}$  (37 mg/ml in dimethyl-sulfoxide) for 60 minutes at 37°C. The reaction is quenched by the addition of 1 ml of water and loaded on a C18 Sep-Pak cartridge (Waters), and the column is washed twice with 5 ml of water to remove unreacted UDP-[ $^3$ H]galactose. The [ $^3$ H]galactosylated  $\text{GlcNAc}\beta\text{O}-(\text{CH}_2)_4-\text{CO}_2\text{Me}$  remains bound to the column during the water washes and is eluted with 5 ml of methanol. Radioactivity in the eluted material is measured by liquid scintillation counting and is proportional to galactosyltransferase activity in the starting sample.

Alternatively, methyltransferase activity is determined using a method that measures transfer of radiolabeled methyl groups from a donor substrate to an acceptor substrate (Bokar, J.A. et al. (*supra*)). Reaction mixtures (50  $\mu$ l final volume) contain 15 mM HEPES, pH 7.9, 1.5 mM  $\text{MgCl}_2$ , 10 mM dithiothreitol, 3% polyvinylalcohol, 1.5  $\mu$ Ci [*methyl*- $^3$ H]AdoMet (0.375  $\mu$ M AdoMet) (DuPont-NEN),

0.6 µg HTFS, and acceptor substrate (0.4 µg [<sup>35</sup>S]RNA or 6-mercaptopurine (6-MP) to 1 mM final concentration). Reaction mixtures are incubated at 30°C for 30 minutes, then 65°C for 5 minutes.

Analysis of [*methyl*-<sup>3</sup>H]RNA is as follows: 1) 50 µl of 2 x loading buffer (20 mM tris-HCl, pH 7.6, 1 M LiCl, 1 mM EDTA, 1% sodium dodecyl sulphate (SDS)) and 50 µl oligo d(T)-cellulose (10 mg/ml in 1 x loading buffer) are added to the reaction mixture, and incubated at ambient temperature with shaking for 30 minutes. 2) Reaction mixtures are transferred to a 96-well filtration plate attached to a vacuum apparatus. 3) Each sample is washed sequentially with three 2.4 ml aliquots of 1 x oligo d(T) loading buffer containing 0.5% SDS, 0.1% SDS, or no SDS. and 4) RNA is eluted with 300 µl of water into a 96-well collection plate, transferred to scintillation vials containing liquid scintillant, and radioactivity determined.

Analysis of [*methyl*-<sup>3</sup>H]6-MP is as follows: 1) 500 µl 0.5 M borate buffer, pH 10.0, and then 2.5 ml of 20% (v/v) isoamyl alcohol in toluene are added to the reaction mixtures. 2) The samples mixed by vigorous vortexing for ten seconds. 3) After centrifugation at 700g for 10 minutes, 1.5 ml of the organic phase is transferred to scintillation vials containing 0.5 ml absolute ethanol and liquid scintillant, and radioactivity determined. and 4) Results are corrected for the extraction of 6-MP into the organic phase (approximately 41%).

## XII. Functional Assays

HTFS function is assessed by expressing the sequences encoding HTFS at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 µg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 µg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in

expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

- 5       The influence of HTFS on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HTFS and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY).
- 10      mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HTFS and other genes of interest can be analyzed by northern analysis or microarray techniques.

### **XIII. Production of HTFS Specific Antibodies**

- HTFS substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g.,
- 15      Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

- Alternatively, the HTFS amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for
- 20      selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

- Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO), by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase
- 25      immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-HTFS activity by, for example, binding the peptide or HTFS to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

### **XIV. Purification of Naturally Occurring HTFS Using Specific Antibodies**

- 30      Naturally occurring or recombinant HTFS is substantially purified by immunoaffinity chromatography using antibodies specific for HTFS. An immunoaffinity column is constructed by covalently coupling anti-HTFS antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.



Media containing HTFS are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HTFS (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HTFS binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HTFS is collected.

#### **XV. Identification of Molecules Which Interact with HTFS**

HTFS, or biologically active fragments thereof, are labeled with  $^{125}\text{I}$  Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HTFS, washed, and any wells with labeled HTFS complex are assayed. Data obtained using different concentrations of HTFS are used to calculate values for the number, affinity, and association of HTFS with the candidate molecules.

Alternatively, molecules interacting with HTFS are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, *Nature* 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

HTFS may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

20

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments.

25

Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	43	016233	HUVELPB01	016233H1 (HUVELPB01), 016233X2 (HUVELPB01), 1594237F1 (BRAINT14), 1610803F6 (COLNTUT06), 1610803H1 (COLNTUT06),
2	44	078336	SYNORAB01	078336F1 (SYNORAB01), 078336H1 (SYNORAB01), 078336R1 (SYNORAB01),
3	45	130117	TESTNOT01	130117H1 (TESTNOT01), 130117R6 (TESTNOT01), 873599R1 (LUNGAST01), 873599T1 (LUNGAST01), 3115426H1 (BRSTNOT17), 3673668H1 (PLACNOT07), SBXA05022D1,
4	46	267495	HNT2NOT01	267495H1 (HNT2NOT01), 267495R6 (HNT2NOT01), 1644251F6 (HEARFET01), 2284760R6 (BRAINON01), 2284760T6 (BRAINON01), 2959833H1 (ADRENOT09), 3040044H1 (BRSTNOT16),
5	47	410533	BRSTNOT01	410533F1 (BRSTNOT01), 410533H1 (BRSTNOT01), 550006R1 (BEPINOT01), 769358T6 (COLNCRT01), 1215984R1 (BRSTTUT01), 1856825T6 (PROSNOT18), 2512020H1 (CONUTUT01), g3095236
6	48	852708	NGANNOT01	852708H1 (NGANNOT01), 915483R1 (BRSTNOT04), 1549163H1 (PROSNOT06), 1618977X17R1 (BRAITUT12), 1968356R6 (BRSTNOT04),
7	49	972944	MUSCNOT02	972944H1 (MUSCNOT02), 1339880F1 (COLNTUT03), 1398783T1 (BRAITUT08), 1534517F1 (SPLNNOT04), 2193889H1 (THYRTUT03), 2543627H1 (UTRSNOT11), 2617014H2 (GBLANOT01), 2777792H1 (OVARUTUT03),
8	50	997730	KIDNTUT01	529402R6 (BRAINOT03), 997730H1 (KIDNTUT01), 1800104F6 (COLNNOT27), 4507277F6 (OVARTDUT01), 5512391H1 (BRADDIR01),
9	51	1285944	COLNNOT16	161632F1 (ADENINB01), 775071R6 (COLNNOT05), 1285944F6 (COLNNOT16), 1285944H1 (COLNNOT16), 1907959F6 (CONNTUT01), 3243618H1 (BRAINT19), 5058828H1 (COLATMT01),
10	52	1293207	PGANNOT03	959637R1 (BRSTTUT03), 1293207H1 (PGANNOT03), 3119913F6 (LUNGUT13),

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
11	53	1308125	COLNFET02	937858H1 (CERVNOT01), 1308125P6 (COLNFET02), 1308125H1 (COLNFET02),
12	54	1439670	PANCNOT08	264074H1 (HNT2AGT01), 1438620P1 (PANCNOT08), 1439670H1 (PANCNOT08), 2671159F6 (ESOGTUT02), 4970569H1 (KIDEUNC10),
13	55	1444281	THYRNOT03	1439532T1 (PANCNOT08), 1444281F1 (THYRNOT03), 1444281H1 (THYRNOT03), 5203108H1 (STOMNOT08),
14	56	1450140	PLACNOT02	1424388R1 (BEPINON01), 1450140F6 (PLACNOT02), 1450140H1 (PLACNOT02), 2102011H1 (BRAITUT02), 2159622H1 (ENDCNOT02), 2187314F6 (PROSNOT26), 2288117H1 (BRAINON01), 3326995H1 (HEAONOT04), 3504033H1 (ADRENOT11),
15	57	1604828	LUNGNOT15	078714F1 (SYNORAB01), 476529R6 (MMLR2DT01), 791628R6 (PROSTUT03), 1604828H1 (LUNGNOT15), 2849341H1 (BRSTTUT13),
16	58	1644023	HEARFET01	1644023CT1 (HEARFET01), 1644023H1 (HEARFET01), 1820555F6 (GBLATUT01), 1930930F6 (COLNTUT03),
17	59	1723402	BLADNOT06	1358516F1 (LUNGNOT09), 1378995F1 (LUNGNOT10), 1476828F1 (CORPNOT02), 1723402H1 (BLADNOT06), 4380622H1 (LUNGNOT37),
18	60	1740585	HIPONON01	1740585H1 (HIPONON01), 2816878F6 (BRSTNOT14), 2816878T6 (BRSTNOT14), 4511074H1 (EPIMNOT01),
19	61	1810925	PROSTUT12	087350R6 (LIVRNOT01), 1810925F6 (PROSTUT12), 1810925H1 (PROSTUT12), 1811144F6 (PROSTUT12), 1928529R6 (BRSTNOT02), 2764684H1 (BRSTNOT12), 4642657H1 (PROSTMT03), SXAF04161V1, SXAF03292V1,
20	62	1915064	PROSTUT04	726814R1 (SYNOOAT01), 1404593H1 (LATRTUT02), 1496266T1 (PROSNON01), 1533375F6 (SPLNNOT04), 1915064H1 (PROSTUT04), 2127869H1 (KIDNNOT05), 2722924H1 (LUNGUT10),
21	63	2185608	PROSNOT26	762593F1 (BRAITUT02), 1393863F1 (THYRNOT03), 2070989X15C1 (ISLTNOT01), 2185608H1 (PROSNOT26), 2449814X13D2 (ENDANOT01),

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
22	64	2228862	PROSNOT16	1510361F6 (LUNGNOT14), 1510361T6 (LUNGNOT14), 1812107F6 (PROSTUT12), 2228862H1 (PROSNOT16), 2718666F6 (THYRNOT09), 2989512H1 (KIDNFET02), 2995876H1 (OVARUT07), 3520458R6 (LUNGNON03), 3594956H1 (FIBPNOT01), 4071960H1 (KIDNNOT26), 4871845H1 (COLDNOT01),
23	65	2235577	PANCTUT02	1570357F1 (UTRSNOT05), 2157703F6 (BRAINOT09), 2235577H1 (PANCTUT02), SBFA00645F1, SBFA02725F1,
24	66	2271680	PROSNON01	223521R1 (PANCTUT01), 1726945T6 (PROSNOT14), 2271680H1 (PROSNON01), 2660011F6 (LUNGNOT09), 2863738H1 (KIDNNOT20), 2866674F6 (KIDNNOT20), 3142517F6 (SMCCNOT02),
25	67	2325603	OVARNOT02	320787R6 (EOSIHET02), 1594544F1 (BRAINOT14), 1597366F6 (BRAINOT14), 2325603H1 (OVARNOT02),
26	68	2356055	LUNGNOT20	2356055H1 (LUNGNOT20), 2356055T6 (LUNGNOT20), 3390020H1 (LUNGNOT17), SBFA00796F1, SBFA00170F1,
27	69	2448909	ENDANOT01	2448909H1 (ENDANOT01), 2615665F6 (GBLANOT01), 2937323F6 (THYMFET02), g1241402
28	70	2631212	COLNTUT15	159493H1 (ADENINB01), 1439529F6 (PANCTUT08), 1626907F6 (COLNPOT01), 1746019H1 (STOMTUT02), 2631212F6 (COLNTUT15), 2631212H1 (COLNTUT15), 2671521H1 (ESOGTUT02), 2871743H1 (THYRNOT10), 2909832H1 (KIDNTUT15), 3077960H1 (BONEUNT01), 3689194H1 (HEAANOT01), 4583971H1 (OVARNOT13), 5674918H1 (293TF2T01),
29	71	2678733	KIDNFET02	434046R6 (THYRNOT01), 2678733H1 (KIDNFET02), 3536993H1 (KIDNNOT25), 3585231H1 (293TF4T01), 3843364H1 (DENDNOT01), 4729211H1 (GBLADIT01), 5535276H1 (HEARFET05),
30	72	2768571	COLANOT02	609405R6 (COLNNOT01), 782088R6 (MYOMNOT01), 2598685F6 (UTRSNOT10), 2768571F6 (COLANOT02), 2768571H1 (COLANOT02), 5058513H1 (COLATMT01), SAQA02758F1, SAQA03638F1, SAQA00977F1,

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
31	73	3189062	THYMN04	2018279F6 (THP1NOT01), 2759759H1 (THP1AZS08), 2777307T6 (OVARUT03), 3189062H1 (THYMN04), 3269328F6 (BRAINOT20), 4289573H1 (BRABDI01), 5275252F6 (OVARIN02),
32	74	3243884	BRAINOT19	266876T6 (HNT2NOT01), 606598H1 (BRSTTUT01), 2203588X49C1 (SPLNFET02), 2837749T6 (DRGLNOT01), 3037267F6 (SMCCNOT02), 3243884H1 (BRAINOT19), 4009730H1 (MUSCNOT10),
33	75	3400578	UTRSNOT16	966455T6 (BRSTNOT05), 2049034T6 (LIVRFET02), 2640820H1 (LUNGUT08), 3400578CT1 (UTRSNOT16), 3400578H1 (UTRSNOT16),
34	76	3422577	UCMCNOT04	019009X308V1 (HUELPEB01), 019009X312D1 (HUELPEB01), 934792H1 (CERVNOT01), 1493507R6 (PROSNON01), 1554828H1 (BLADTUT04), 2751872H1 (THP1AZS08), 3422577H1 (UCMCNOT04), 5019159H1 (PANCNOT22),
35	77	3706809	PENCNOT07	292348H1 (TMLR3DT01), 1319573T1 (BLADNOT04), 1444394F6 (THYRN03), 1456216F1 (COLNFET02), 1456216R1 (COLNFET02), 2414862F6 (HNT3AZT01), 2717052H1 (THYRN09), 3706809H1 (PENCNOT07),
36	78	3745914	THYMN08	440059R6 (THYRN01), 463957R6 (LATRNOT01), 3745914H1 (THYMN08), SBOA01668D1, SBOA00128D1,
37	79	4000776	HNT2AZS07	2888759F6 (LUNGFET04), 4000776H1 (HNT2AZS07), 4000776R6 (HNT2AZS07), 4000776T6 (HNT2AZS07),
38	80	4071304	KIDNNOT26	1427681F1 (SINTBST01), 4071304H1 (KIDNNOT26), 4333229H1 (KIDCTMT01), 5342286H1 (CONFNOT05), 5390966H1 (KIDNNOT32), g1102520
39	81	4344970	LYMBTXX01	1215288R6 (BRSTTUT01), 1749984T6 (STOMTUT02), 4187556H1 (BRSTNOT31), 4344970H1 (LYMBTXX01),
40	82	5392302	KIDNNOT32	1748734T6 (STOMTUT02), 2703005F6 (OVARUT10), 3697713H1 (SININOT05), 5392302H1 (KIDNNOT32),

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
41	83	5555235	TONSDIT01	1503473T1 (BRAITUT07), 3199747T6 (PENCNOT02), 4556633H1 (KERAUNT01), 5208015H1 (BRAFN0T02), 5555235H1 (TONSDIT01),
42	84	5573296	TYMNOT08	130117R6 (TESTNOT01), 873599R1 (LUNGAST01), 1356271F6 (LUNGNOT09), 1844333T6 (COLNNOT08), 3115426H1 (BRSTNOT17), 3673668H1 (PLACNOT07), 4339255H1 (BRAUNOT02), 5573296H1 (TYMNOT08), SBXA05492D1, SBXA02308D1,

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
1	261	T238 T97 S118 T157 S31 T97 S118 S164 S172 T200 S234 S248 S257		Phosphoribulokinase profile: D147-P199 Phosphoribulokinase family signatures: V131-F149 R166-S183 Y185-I208 Phosphoribulokinase (kinase/transferase) region: L22-G232 ATP/GTP binding site motif A (P-loop): G27-S34	g471981 uridine kinase [Mus musculus]	BLAST-GenBank MOTIFS PROFILES-SCAN BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
2	197	S37	N82		g2104536 predicted glycosyl transferase [Arabidopsis thaliana]	BLAST-GenBank MOTIFS
3	378	S69 S107 T293 T359 S109 S145 T181 T214 S217		Transmembrane domain: L336-M355 Hypothetical transmembrane protein region: Y169-T297 K43-L162 Acyltransferase-like region: V20-Y320	g8886001 lysophosphatidic acid acyltransferase [Homo sapiens]	BLAST-GenBank MOTIFS HMMER BLAST-PRODOM BLAST-DOMO

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
4	285	S27 T90 S144 S271 T16 T139 T236 Y49 Y187 Y187	N110 N275	Uroporphyrin-III C-methyltransferase signature: I5-L31	g171416 DPH5 (Diphthine synthase/Diphthamide biosynthesis methyltransferase) (EC 2.1.1.98) [Saccharomyces cerevisiae]	BLAST-GenBank MOTIFS BLIMPS-BLOCKS BLAST-PRODOM BLAST-DOMO
5	301	S101 S166 T185 S26 T219 T233		CDP-alcohol phosphatidyl transferase profile: F133-A187 CDP-, alcohol phosphatidyl transferase-like region: W107-A207	g2414601 phosphatidyl synthase (transferase) [Schizosaccharomyces pombe]	BLAST-GenBank MOTIFS PROFILES-SCAN BLAST-DOMO
6	253	T3 S12 T63 T96 S135 T108 T122 T137		SH3 domain: E33-H87 Arginine N-methyltransferase 2-like regions (EC 2.1.1.-): G85-T137 M1-E34	g1655625 arginine methyltransferase [Homo sapiens]	BLAST-GenBank MOTIFS HMMER-PFAM BLAST-PRODOM
7	390	T43 S141 S153 S370 S2 S17 T54 S252	N251	S-malonyltransferase-like region: L66-K359	g5139348 S-malonyltransferase (acyl-carrier protein) [Brassica napus]	BLAST-GenBank MOTIFS BLAST-DOMO
8	373	S101 S292 S6 T27 T119 T120 T149 T276	N114 N203 N282 N328	Signal peptide: M1-G23 Lipoate ligase-like region: I33-I307	g2780412 lipoyltransferase [Bos taurus]	BLAST-GenBank MOTIFS BLAST-DOMO HMMER



Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
9	371	S233 S317 T365 S297 S341 S2 S117 S120 S146 S163 S204 S233 T355	N103 N249 N257	Signal peptide: M1-H23	g2104536 predicted glycosyl transferase [Arabidopsis thaliana]	BLAST-GenBank MOTIFS HMMER SPSCAN
10	123	S32 S94		S-adenosyl methionine synthetase (Methionine adenosyl transferase) profile: N49-L117		MOTIFS PROFILESCAN
11	85	S4 S41 T63 T73 S3 S68			g2511715 putative phosphatidylinositol-4-phosphate 5-kinase [Arabidopsis thaliana] P=2.1e-05	BLAST-G nBank MOTIFS
12	184	T87 T171 S145			g7406641 glucosamine-6-phosphate acetyltransferase (EMeg32) [Mus musculus]. Boehmelt, G. et al. (2000) J. Biol. Chem. 275:12821-12832.	BLAST-G nBank MOTIFS
13	169	T20 T70 T59 S96 T150	N18 N82		g9998952 ethanolamine kinase [Homo sapiens]	BLAST-GenBank MOTIFS

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
14	357	S6 S223 S302 T26 T254 S346		Protein-L-isoaspartate (D-aspartate) O-methyltransferase signatures: C67-H114 (score: strength=0.73) G148-I191 (score: strength=0.60) Protein-L-isoaspartate (D-aspartate) O-methyltransferase -like region: L14-L219	g2621917 L-isoaspartyl protein carboxyl methyltransferase	BLAST-GenBank MOTIFS BLIMPS-BLOCKS BLAST-PRODOM
15	100	T87 S27	N26	Signal peptide: M1-W22 or M1-S27 Transmembrane domain: V7-F21 Alpha 13, galactosyltransferase-like region: M1-R84	g7717225 alpha 1,3 galactosyltransferase [Platyrrhini]. Henion, T.R. et al. (1994) Glycobiology 4:193-201.	BLAST-GenBank MOTIFS BLAST-DOMO BLAST-PRODOM HMMER SPSCAN
16	199	S130 T131 S34 S43 T84 S88 T171 S173 S14 T47 S76 S130	N29 N113	Uridine kinase signatures: T3-A20 S31-E42 ATP/GTP binding site motif A (P-loop): G10-T17	g6224931 uridine kinase [Danio rerio].	BLAST-GenBank MOTIFS BLIMPS-PRINTS

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
17	244	S73 T95 S110 S53 S69 S149 S208	N203	Signal peptide: M1-S29 Transmembrane domain: I7-G26 ubiE/COQ5 methyltransferase signature: Q132-V176 (score: strength=0.67)	G5458322 ubiquinone/menaquinone biosynthesis methyl transferase (ubiE) [Pyrococcus abyssi].	BLAST-GenBank MOTIFS BLIMPS-BLOCKS HMMER
18	358	S171 T10 T245 T287 S158 T263 T279 T287	N105	Signal peptide: M1-S32 Polynucleotide kinase/5' hydroxyl transferase-like regions: (P=1.3e-04) L20-M59, Y135-K211 ATP/GTP binding site motif A (P-loop): G25-S32		MOTIFS HMMER BLAST-PRODOM
19	302	S20 S48 S109 S151 S265 S7 S24 S122 S144 S273	N5 N142		G2828262 aralkyl acyl-CoA:amino acid N-acyltransferase [Bos taurus]	BLAST-GenBank MOTIFS
20	234	S103 S153 T202 S6 T61 S79 S103 S130 T137			G4200446 FYVE finger-containing phosphoinositide kinase [Mus musculus]. Shisheva, A. et al. (1999) Mol. Cell. Biol. 19:623-634.	BLAST-GenBank MOTIFS

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
21	403	S116 S129 T212 T218 S231 S363 T53 S178 T218 T247 T286 Y176 Y257	N134	Glycosyl transferase/queuine tRNA-ribosyltransferase-like region: M17-M377	g7415808 tRNA-guanine transglycosylase [Mus musculus].	BLAST-GenBank MOTIFS BLAST-PRODOM
22	487	T47 S24 T47 T80 T85 T146 T389 T93 S269 T272 T283 S310 T437 Y52 Y52	N181 N237	Transmembrane domains: W314-M340 F127-L143 Phosphatidylserine synthase I (transmembrane transferase)-like region: R44-W429	g2190007 phosphatidylserine synthase II [Cricetulus griseus]	BLAST-GenBank MOTIFS HMMER BLAST-PRODOM
23	246	S66 T142 S212 S206	N64 N140	ubiE/COQ5 methyltransferase signatures: P123-S167 V185-D196 L203-K245 Ubiquinone methyltransferase-like region: I134-K245	g8777363 ubiquinone/menaquinone biosynthesis methyltransferase-like [Arabidopsis thaliana]. Sato, S. et al. (2000) DNA Res. 7:31-63.	BLAST-GenBank MOTIFS BLIMPS-BLOCKS BLAST-PRODOM
24	410	S108 S114 T131 S160 S245 S333 S3 S63 S76	N295 N362	Signal peptide: M1-T25 Methyltransferase/demethylubiquinone-like region: V133-H370	g7673361 methyltransferase COQ3 [Homo sapiens]. Jonassen, T. and Clarke, C.F. (2000) J. Biol. Chem. 275:12381-12387.	BLAST-GenBank MOTIFS HMMER BLAST-PRODOM

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
25	253	S71 S171 S244 S16 T81 T223 S240 S244	N2		g2414623 putative phosphotransferase [Schizosaccharomyces pombe]	BLAST-GenBank MOTIFS
26	303	T6 T58 S77 T82 S248 S253 T87 S121 T198 S250 T257		NNMT/PNMT/TEMT methyltransferase family signatures: Q12-L38, L42-D85 F86-W107, A110-A134 L156-L199 NNMT/PNMT/TEMT methyltransferase family domain: M1-K233	g6580815 indolethylamine N-methyltransferase [Homo sapiens]. Thompson, M.A. et al. (1999) Genomics 61:285-297.	BLAST-GenBank MOTIFS
27	307	S67 S76 S172 T8 S28 T58 S116 S154 S172 T259	N26		g6466950 putative ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit N-methyltransferase I [Arabidopsis thaliana].	BLAST-GenBank MOTIFS
28	169	S4 S30 S119 T134	N117	Acetyltransferase (GNAT) signatures (E=0.01): C79-G89 (Score=976) A120-F129 Acetyltransferase-like region: D33-E130	g7688322 putative GNAT family acetyltransferase with 2 zinc fingers [Schizosaccharomyces pombe].	BLAST-GenBank MOTIFS BLIMPS-PFAM BLAST-PRODOM
29	389	T200 T341 S90 T282 T369 T374	N292	Phosphoribosylglycinamide formyltransferase-like region: V50-L216	g3288685 mitochondrial methionyl-tRNA transformylase [Bos taurus]	BLAST-GenBank MOTIFS BLAST-DOMO

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
30	600	T74 T175 T96 T179 T278 S313 S138 S148 S155 T171 S181 T191 S211 T245 S282 S359 S557 T595 Y268 Y554	N300 N311 N331 N375 N460	Signal peptide: M1-A30 Sialyltransferase family signatures: I363-S417 G511-T556 Golgi transferase/alpha- N-acetylgalactosaminide alpha-2,6- sialyltransferase-like region (E.C.2.4.99.3): G426-R590	g6491775 Gainac alpha-2,6- sialyltransferase I [Mus musculus]. Lee, Y.C. et al. (1999) J. Biol. Chem. 274:11958- 11967.	BLAST-GenBank MOTIFS HMME BLIMPS-PFAM BLAST-PRODOM
31	448	S8 T65 S313 S353 T388 S14 T30 S95 T223	N221	Precorrin-6 methyltransferase-like region (P=9.7e-09): E224-G309		BLAST-GenBank MOTIFS BLAST-PRODOM
32	346	S324 T10 S152 S295 S324 T2 T10 T37 T107 S117 S152 T270 T282	N46 N150	Ribosomal RNA adenine dimethylase domain: Q35-D305 Ribosomal RNA adenine dimethylase signatures: T44-R89 I138-I151 I212-L233	g2073482 dimethyladenosine transferase [Rickettsia prowazekii]	BLAST-GenBank MOTIFS BLIMPS-BLOCKS HMME-PFAM BLAST-DOMO BLAST-PRODOM

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
33	173	S38 S101	N113	Signal peptide: M1-A24 Nucleoside diphosphate kinase domain: K76-E143 Nucleoside diphosphate kinases-active site profile: A94-W140 Nucleoside diphosphate kinase (transferase)-like region: L11-D134	g3228530 type 6 nucleoside diphosphate kinase NM23-H6 [Homo sapiens]	BLAST-GenBank MOTIFS HMMER HMMER-PFAM PROFILESCAN BLAST-PRODOM BLAST-DOMO
34	445	S21 T31 S33 S37 S47 T144 S440 T441 S63 S255 T329 T415 T427	N61 N295	Transmembrane domains: L77-C100 F168-V186 I223-F242 Phosphatidate cytidylyltransferase domain: L59-I401 Phosphatidate cytidylyltransferase signatures: G221-I252 K253-G266 L342-C385 Cytidylyltransferase-like region: W124-I401	g1915972 CDP-diacylglycerol synthase [Homo sapiens]	BLAST-GenBank MOTIFS HMMER HMMER-PFAM BLIMPS-BLOCKS BLAST-DOMO BLAST-PRODOM

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
35	420	S105 S66 S85 T86 S147 S299		Signal peptide: M1-R28 SpoU rRNA methylase family signature (P<0.011): P220-A230 I380-E401	g7226378 RNA methyltransferase, TrmH family [Neisseria meningitidis MC58]. Tettelin, H. et al. (2000) Science 287:1809-1815.	BLAST-GenBank MOTIFS HMMER BLIMPS-PFAM
36	354	T146 T70 T80 T176 T257 T258 S111 T146 T176 S203 S223 S224 S297 Y84 Y213	N256 N286	Signal peptide: M1-S26 M1-A30 UDP-glucose:glycoprotein glucosyltransferase precursor-like region (E.C.2.4.1.-): S26-G354	g7677176 UDP-glucose glycoprotein:glucosyltransferase precursor [Rattus norvegicus]. C Tessier, D. et al. (2000) Glycobiology 10:403-412.	BLAST-GenBank MOTIFS HMMER SPSCAN BLAST-PRODOM
37	198	S6 T147 S160 T165 S183 S190 S2 S3 S51 T55 S128 S152 T165 S110 T139 S171 T233 S263 S17 S25 S123 T129 T139 S193		Phosphoribosylglycinamide formyltransferase active site profile: S48-G105		MOTIFS PROFILES SCAN
38	296				g2828262 aralkyl acyl-CoA:amino acid N-acyltransferase [Bos taurus]	BLAST-G nBank MOTIFS
39	214	T173 T89 T113 S164 T173		N-6 adenine-specific DNA methylase signature (N-terminal region) and motif: V100-A143, L119-Y125	g5052364 putative N6-DNA-methyltransferase; N6AMT1 [Homo sapiens]	BLAST-GenBank MOTIFS PROFILES SCAN BLIMPS-PRINTS



Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification	Analytical Methods
40	322	T159 T191 S201 S213 S245 T247 S286 T32 T256	N199 N284 N292	pfkB family of carbohydrate kinases domain: G52-A274 pfkB family of carbohydrate kinases signatures: F51-G66 G228-G240 V262-L275 pfkB family of carbohydrate kinases profiles: T32-F97 C241-S301 Ribokinase signatures: C20-T41 F47-G66 N119-N132 Q188-A203 V232-V243	G4959404 ribokinase [Lactobacillus sakei]	BLAST-GenBank MOTIFS PROFILES BLIMPS-PRINTS BLIMPS-BLOCKS BLAST-DOMO BLAST-PRODOM
41	87	T12		Rhodanese (thiosulfate sulfurtransferase) (E.C.2.8.1.1) C-terminal profile: G33-H86		MOTIFS PROFILES
42	378	S69 S107 T293 T359 S109 S145 T181 T214 S217		Transmembrane domain: L336-M355 Acyltransferase-like region: V20-Y320	G8886001 lysophosphatidic acid acyltransferase- gamma [Homo sapiens].	BLAST-GenBank MOTIFS HMMER BLAST-DOMO

Table 3

SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
43	Reproductive (0.286) Gastrointestinal (0.250) Hematopoietic/Immune (0.179)	Cancer (0.500) Cell Proliferation (0.321) Inflammation (0.214)	PBLUESCRIPT
44	Nervous (0.276) Reproductive (0.276) Cardiovascular (0.172)	Cancer (0.448) Cell Proliferation (0.207) Inflammation (0.207)	PBLUESCRIPT
45	Reproductive (0.269) Musculoskeletal (0.192) Nervous (0.154)	Cancer (0.346) Inflammation (0.269) Cell Proliferation (0.269)	PBLUESCRIPT
46	Reproductive (0.269) Gastrointestinal (0.135) Hematopoietic/Immune (0.135)	Cancer (0.462) Cell Proliferation (0.365) Inflammation (0.173)	PBLUESCRIPT
47	Reproductive (0.317) Gastrointestinal (0.183) Cardiovascular (0.167)	Cancer (0.633) Cell Proliferation (0.200) Inflammation (0.200)	PBLUESCRIPT
48	Reproductive (0.252) Nervous (0.175) Cardiovascular (0.175)	Cancer (0.437) Cell Proliferation (0.223) Inflammation (0.223)	PSPORT1
49	Nervous (0.250) Reproductive (0.156) Endocrine (0.156)	Cancer (0.469) Inflammation (0.250) Cell Proliferation (0.125)	PSPORT1
50	Reproductive (0.278) Nervous (0.222) Cardiovascular (0.222)	Cancer (0.556) Inflammation (0.389) Trauma (0.167)	PSPORT1
51	Reproductive (0.238) Nervous (0.214) Gastrointestinal (0.143)	Cancer (0.488) Inflammation (0.238) Cell Proliferation (0.179)	pINCY
52	Reproductive (0.302) Gastrointestinal (0.151) Cardiovascular (0.132)	Cancer (0.472) Cell Proliferation (0.226) Inflammation (0.132)	pINCY
53	Reproductive (0.500) Gastrointestinal (0.250) Developmental (0.250)	Inflammation (0.500) Cancer (0.250) Cell Proliferation (0.250)	pINCY

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
54	Gastrointestinal (0.375) Reproductive (0.250) Cardiovascular (0.188)	Cancer (0.625) Cell Proliferation (0.250) Trauma (0.188)	pINCY
55	Endocrine (0.250) Gastrointestinal (0.250) Hematopoietic/Immune (0.250) Urologic (0.250)	Cancer (0.750) Inflammation (0.250)	pINCY
56	Reproductive (0.286) Nervous (0.143) Cardiovascular (0.179)	Cancer (0.500) Inflammation (0.179) Cell Proliferation (0.179)	pINCY
57	Reproductive (0.235) Musculoskeletal (0.235) Cardiovascular (0.176)	Cancer (0.588) Inflammation (0.294)	pINCY
58	Reproductive (0.269) Gastrointestinal (0.135) Cardiovascular (0.135)	Cancer (0.558) Cell Proliferation (0.135) Trauma (0.115) Inflammation (0.115)	pINCY
59	Reproductive (0.245) Nervous (0.286) Cardiovascular (0.143)	Cancer (0.490) Inflammation (0.224) Cell Proliferation (0.122)	pINCY
60	Reproductive (0.385) Nervous (0.231) Cardiovascular (0.154)	Cancer (0.538) Inflammation (0.154) Trauma (0.231)	PSPORT1
61	Reproductive (0.474) Urologic (0.263) Gastrointestinal (0.211)	Cancer (0.579) Inflammation (0.158) Cell Proliferation (0.158)	pINCY
62	Reproductive (0.270) Nervous (0.140) Gastrointestinal (0.160)	Cancer (0.472) Inflammation (0.160) Cell Proliferation (0.151)	PSPORT1
63	Reproductive (0.245) Nervous (0.286)	Cancer (0.510) Inflammation (0.180) Cell Proliferation (0.130)	pINCY

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
64	Reproductive (0.341) Nervous (0.136) Cardiovascular (0.136)	Cancer (0.591) Inflammation (0.114) Cell Proliferation (0.250)	pINCY
65	Reproductive (0.310) Nervous (0.190) Cardiovascular (0.121) Hematopoietic/Immune (0.121)	Cancer (0.483) Inflammation (0.276) Cell Proliferation (0.155)	pINCY
66	Reproductive (0.304) Nervous (0.174) Cardiovascular (0.130) Urologic (0.130)	Cancer (0.522) Inflammation (0.174) Cell Proliferation (0.261)	PSPORT1
67	Reproductive (0.169) Nervous (0.271) Hematopoietic/Immune (0.203)	Cancer (0.339) Inflammation (0.322) Cell Proliferation (0.169)	PSPORT1
68	Cardiovascular (0.429) Reproductive (0.286) Nervous (0.143) Developmental (0.143)	Cancer (0.571) Cell Proliferation (0.143) Trauma (0.143)	pINCY
69	Developmental (0.375) Cardiovascular (0.250) Hematopoietic/Immune (0.125) Reproductive (0.250)	Cancer (0.375) Inflammation (0.125) Cell Proliferation (0.625)	PBLUESCRIPT
70	Reproductive (0.198) Nervous (0.148) Gastrointestinal (0.210)	Cancer (0.494) Inflammation (0.247) Cell Proliferation (0.185)	pINCY
71	Urologic (0.182) Nervous (0.182) Endocrine (0.182)	Cancer (0.273) Inflammation (0.273) Cell Proliferation (0.273)	pINCY
72	Gastrointestinal (0.500) Reproductive (0.367) Nervous (0.100)	Cancer (0.567) Inflammation (0.233) Trauma (0.100)	pINCY

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
73	Reproductive (0.333) Nervous (0.143) Hematopoietic/Immune (0.238)	Cancer (0.476) Inflammation (0.143) Cell Proliferation (0.190) Trauma (0.143)	PSPORT1
74	Reproductive (0.300) Nervous (0.200) Cardiovascular (0.133)	Cancer (0.467) Inflammation (0.300) Cell Proliferation (0.233)	pINCY
75	Reproductive (0.250) Gastrointestinal (0.250) Cardiovascular (0.167)	Cancer (0.458) Inflammation (0.250) Cell Proliferation (0.292)	pINCY
76	Reproductive (0.333) Nervous (0.140) Cardiovascular (0.140)	Cancer (0.544) Inflammation (0.193) Cell Proliferation (0.158)	pINCY
77	Reproductive (0.184) Nervous (0.211) Gastrointestinal (0.132) Hematopoietic/Immune (0.132) Urologic (0.132)	Cancer (0.447) Inflammation (0.237) Cell Proliferation (0.184)	pINCY
78	Reproductive (0.167) Nervous (0.333) Cardiovascular (0.333) Endocrine (0.167)	Cancer (0.333) Inflammation (0.333) Cell Proliferation (0.167)	pINCY
79	Nervous (0.500) Developmental (0.500)	Cell Proliferation (1.000)	PSPORT1
80	Urologic (0.500) Nervous (0.250) Gastrointestinal (0.250)	Cancer (0.750) Inflammation (0.250)	pINCY
81	Reproductive (0.400) Musculoskeletal (0.133) Cardiovascular (0.133)	Cancer (0.667) Inflammation (0.067)	pINCY
82	Gastrointestinal (0.318) Hematopoietic/Immune (0.227) Cardiovascular (0.182)	Cancer (0.455) Inflammation (0.364) Cell Proliferation (0.182)	pINCY

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
83	Reproductive (0.227) Nervous (0.170) Cardiovascular (0.182)	Cancer (0.500) Inflammation (0.239) Cell Proliferation (0.205)	pINCY
84	Reproductive (0.269) Nervous (0.154) Musculoskeletal (0.192)	Cancer (0.346) Inflammation (0.269) Cell Proliferation (0.269)	pINCY

Table 4

SEQ ID No:	Library	Library Comment
43	HUVELPB01	Library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells that were stimulated with cytokine/LPS. RNA was isolated from two pools of HUV-EC-C cells that had been treated with either gamma IFN and TNF-alpha or IL-1 beta and LPS. In the first instance, HUV-EC-C cells were treated with 4 units/ml TNF and 2 units/ml IFNg for 96 hours. In the second instance, cells were treated with 1 units/ml IL-1 and 100 ng/ml LPS for 5 hours.
44	SYNORAB01	Library was constructed using RNA isolated from the synovial membrane tissue of a 68-year-old Caucasian female with rheumatoid arthritis.
45	TESTNOT01	Library was constructed using RNA isolated from the testicular tissue of a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
46	HNT2NOT01	Library was constructed at Stratagene (STR937230), using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor).
47	BRSTNOT01	Library was constructed using RNA isolated from the breast tissue of a 56-year-old Caucasian female who died in a motor vehicle accident.
48	NGANNOT01	Library was constructed using RNA isolated from tumorous neuroganglion tissue removed from a 9-year-old Caucasian male during a soft tissue excision of the chest wall. Pathology indicated a ganglioneuroma. Family history included asthma.
49	MUSCNOT02	Library was constructed using RNA isolated from the psoas muscle tissue of a 12-year-old Caucasian male.
50	KIDNTUT01	Library was constructed using RNA isolated from the kidney tumor tissue removed from an 8-month-old female during nephroureterectomy. Pathology indicated Wilms' tumor (nephroblastoma), which involved 90 percent of the renal parenchyma. Prior to surgery, the patient was receiving heparin anticoagulant therapy.
51	COLNNOT16	Library was constructed using RNA isolated from sigmoid colon tissue removed from a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy.
52	PGANNOT03	Library was constructed using RNA isolated from paraganglionic tumor tissue removed from the intra-abdominal region of a 46-year-old Caucasian male during exploratory laparotomy. Pathology indicated a benign paraganglioma and was associated with a grade 2 renal cell carcinoma, clear cell type, which did not penetrate the capsule. Surgical margins were negative for tumor.
53	COLNFET02	Library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus, who died at 20 weeks' gestation.

Table 4

SEQ ID NO:	Library	Library Comment
54	PANCNOT08	Library was constructed using RNA isolated from pancreatic tissue removed from a 65-year-old Caucasian female during radical subtotal pancreatectomy. Pathology for the associated tumor tissue indicated an invasive grade 2 adenocarcinoma. Patient history included type II diabetes, osteoarthritis, cardiovascular disease, benign neoplasm in the large bowel, and a cataract. Previous surgeries included a total splenectomy, cholecystectomy, and abdominal hysterectomy. Family history included cardiovascular disease, type II diabetes, and stomach cancer.
55	THYRNOT03	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid. >THYRNOT08 PENCY The THYRNOT08 library was constructed isolated from the diseased left thyroid tissue removed from a 13-year-old Caucasian female during a complete thyroidectomy. Pathology indicated lymphocytic thyroiditis.
56	PLACNOT02	Library was constructed using RNA isolated from the placental tissue of a Hispanic female fetus, who was prematurely delivered at 21 weeks' gestation. Serologies of the mother's blood were positive for CMV (cytomegalovirus).
57	LUNGNOT15	Library was constructed using RNA isolated from lung tissue removed from a 69-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated residual grade 3 invasive squamous cell carcinoma. Patient history included acute myocardial infarction, prostatic hyperplasia, and malignant skin neoplasm. Family history included cerebrovascular disease, type I diabetes, acute myocardial infarction, and arteriosclerotic coronary disease.
58	HEARFET01	Library was constructed using RNA isolated from heart tissue removed from a Hispanic male fetus, who died at 18 weeks' gestation.
59	BLADNOT06	Library was constructed using RNA isolated from the posterior wall bladder tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy and urinary diversion. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder and urothelium. Patient history included lung neoplasm, and tobacco abuse in remission. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.



Table 4

SEQ ID NO:	Library	Library Comment
60	HIPONON01	Normalized hippocampus library was constructed from 1.13M independent clones from the HIPONOT01 library. RNA was isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).
61	PROSTUT12	Library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA).
62	PROSTUT04	Library was constructed using RNA isolated from prostate tumor tissue removed from a 57-year-old Caucasian male during radical prostatectomy, removal of both testes and excision of regional lymph nodes. Pathology indicated adenocarcinoma (Gleason grade 3+3). Patient history included a benign neoplasm of the large bowel and type I diabetes. Family history included a malignant neoplasm of the prostate and type I diabetes.
63	PROSNOT26	Library was constructed using RNA isolated from prostate tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology for the matched tumor tissue indicated an adenocarcinoma (Gleason grade 3+4) forming a predominant mass involving the right and left sides anteriorly. The right and left apex and right and left bladder base surgical margins were positive for tumor. The patient presented with elevated PSA. Patient history included benign hypertension. Family history included malignant stomach neoplasm.
64	PROSNOT16	Library was constructed using RNA isolated from diseased prostate tissue removed from a 68-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+4). The patient presented with elevated prostate specific antigen (PSA). During this hospitalization, the patient was diagnosed with myasthenia gravis. Patient history included osteoarthritis, and type II diabetes. Family history included benign hypertension, acute myocardial infarction, hyperlipidemia, and arteriosclerotic coronary artery disease.
65	PANCTUT02	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.

Table 4

SEQ ID NO:	Library	Library Comment
66	PROSNON01	Normalized prostate library was constructed from 4.4 M independent clones from the PROSNOT11 library. Starting RNA was made from prostate tissue removed from a 28-year-old Caucasian male who died from a self-inflicted gunshot wound. The normalization and hybridization conditions were adapted from Soares, M.B. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228-9232, using a longer (19 hour) reannealing hybridization period.
67	OVARNOT02	Library was constructed using RNA isolated from ovarian tissue removed from a 59-year-old Caucasian female who died of a myocardial infarction. Patient history included cardiomyopathy, coronary artery disease, previous myocardial infarctions, hypercholesterolemia, hypotension, and arthritis.
68	LUNGNOT20	Library was constructed using RNA isolated from right upper lobe lung tissue removed from a 61-year-old Caucasian male. Pathology indicated panacinar emphysema with blebs in the right anterior upper lobe and apex, as well as emphysema in the right posterior upper lobe. Patient history included angina pectoris, and gastric ulcer. Family history included a subdural hemorrhage, cancer of an unidentified site, atherosclerotic coronary artery disease, and pneumonia.
69	ENDANOT01	Library was constructed using RNA isolated from aortic endothelial cell tissue from an explanted heart removed from a male during a heart transplant.
70	COLNTUT15	Library was constructed using RNA isolated from colon tumor tissue obtained from a 64-year-old Caucasian female during a right hemicolectomy with ileostomy and bilateral salpingo-oophorectomy (removal of the fallopian tubes and ovaries). Pathology indicated an invasive grade 3 adenocarcinoma. Patient history included hypothyroidism, depression, and anemia. Family history included colon cancer and uterine cancer.
71	KIDNFET02	Library was constructed using RNA isolated from kidney tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart and died at 23 weeks' gestation.
72	COLANOT02	Library was constructed using RNA isolated from diseased ascending colon tissue removed from a 25-year-old Caucasian female during a multiple segmental resection of the large bowel. Pathology indicated moderately to severely active chronic ulcerative colitis, involving the entire colectomy specimen and sparing 2 cm of the attached ileum. Grossly, the specimen showed continuous involvement from the rectum proximally; marked mucosal atrophy and no skip areas were identified. Microscopically, the specimen showed dense, predominantly mucosal inflammation and crypt abscesses. Patient history included benign large bowel neoplasm. Previous surgeries included a polypectomy.

Table 4

SEQ ID NO:	Library	Library Comment
73	THYMNON04	Normalized thymus library was constructed from 1.48 million independent clones from a thymus tissue library. Starting RNA was made from thymus tissue removed from a 3-year-old Caucasian male who died of anoxia.
74	BRAINOT19	Library was constructed using RNA isolated from diseased brain tissue removed from the left frontal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology indicated a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. The left temporal lobe, including the mesial temporal structures, showed focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. GFAP was positive for astrocytes. Family history included brain cancer.
75	UTRSNOT16	Library was constructed using RNA isolated from uterine endometrial tissue removed from a 48-year-old Caucasian female during a vaginal hysterectomy, rectocele repair, and bilateral salpingo-oophorectomy. Pathology indicated chronic cervicitis, and the endometrium was weakly proliferative. The uterus, tubes, ovaries, and specimen from the peritoneum indicated endometriosis focally involving the surface of the right ovary and the peritoneum. Pathology for the associated tumor tissue indicated a single submucosal leiomyoma, which exhibited extensive hyalin change with hyalin-type necrosis. The left ovary contained a corpus luteum cyst. Patient history included hyperlipidemia and meningitis. Family history included benign hypertension, hyperlipidemia, atrial fibrillation, atherosclerotic coronary artery disease, and type II diabetes.
76	UCMCNOT04	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of multiple individuals of mixed age and sex. The cells were treated with G-CSF.
77	PENCNOT07	Library was constructed using RNA isolated from penis right corpora cavernosa tissue removed from a male.
78	THYMNOT08	Library was constructed using RNA isolated from thymus tissue removed from a 4-month-old Caucasian male during a total thymectomy and open heart repair of atrioventricular canal defect using hypothermia. Patient presented with congenital heart anomaly, congestive heart failure, and Down syndrome. Patient history included abnormal thyroid function study and premature birth.

Table 4

SEQ ID NO:	Library	Library Comment
79	HNT2AZS07	Subtracted library was constructed from RNA isolated from an hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor) treated for three days with 0.35 micromolar AZ. The hybridization probe for subtraction was derived from a similarly constructed library from untreated hNT2 cells. 13.08M clones from the AZ-treated library were subjected to three rounds of subtractive hybridization with 3.04M clones from the untreated library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (NAR (1991) 19:1954) and Bonaldo et al. (Genome Research (1996) 6:791).
80	KIDNNOT26	Library was constructed using RNA isolated from left kidney medulla and cortex tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Pathology for the associated tumor tissue indicated grade 2 renal cell carcinoma involving the lower pole of the kidney. Patient history included hyperlipidemia, cardiac dysrhythmia, metrorrhagia, normal delivery, cerebrovascular disease, atherosclerotic coronary artery disease, and tobacco abuse. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
81	LYMBTXX01	Library was constructed using RNA isolated from a treated K-562 cell line, derived from chronic myelogenous leukemia precursor cells removed from a 53-year-old female. The cells were treated with 9cis retinoic acid (RA), 1 micromolar, for 13 days.
82	KIDNNOT32	Library was constructed using RNA isolated from kidney tissue removed from a 49-year-old Caucasian male who died from an intracranial hemorrhage and cerebrovascular accident. Patient history included tobacco abuse.
83	TONSDIT01	Library was constructed using RNA isolated from the tonsil tissue of a 6-year-old Caucasian male during adenotonsillectomy. Pathology indicated lymphoid hyperplasia of the tonsils. Family history included hypothyroidism and benign skin neoplasm.
84	TYMNOT08	Library was constructed using RNA isolated from anergic/allogenic T-lymphocyte tissue removed from an adult (40-50-year-old) Caucasian male. The cells were incubated for 3 days in the presence of OKT3 mAb (1 microgram/ml OKT3) and 5% human serum.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type; FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLOCKS IMPROVED Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score > GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <u>supra</u> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
  - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-42,
  - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-42,
  - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42, and
  - 10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-42.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
- 20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:43-84.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
- 30 9. A method for producing a polypeptide of claim 1, the method comprising:
  - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
  - 35 b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- 5           a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84,  
          b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:43-84,  
          c) a polynucleotide sequence complementary to a),  
          d) a polynucleotide sequence complementary to b), and  
10          e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

15          13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization  
20          complex is formed between said probe and said target polynucleotide or fragments thereof, and  
          b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

25

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and  
30          b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

35



17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-42.

18. A method for treating a disease or condition associated with decreased expression of functional HTFS, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- 10 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

15

21. A method for treating a disease or condition associated with decreased expression of functional HTFS, comprising administering to a patient in need of such treatment a composition of claim 20.

20

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

25

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

30

24. A method for treating a disease or condition associated with overexpression of functional HTFS, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- 35 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- 5           a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound  
10 with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target  
15 polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- 20           c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- 25           b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- 30           c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

## SEQUENCE LISTING

<110> INCYTE GENOMICS, INC.  
 TANG, Y. Tom  
 YUE, Henry  
 HILLMAN, Jennifer L.  
 LAL, Preeti  
 BANDMAN, Olga  
 PATTERSON, Chandra  
 SHIH, Leo  
 AZIMZAI, Yalda  
 LU, Dyung Aina M.  
 BAUGHN, Mariah R.

<120> HUMAN TRANSFERASE MOLECULES

<130> PF-0753 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/163,595

(151) 1999-11-04

<160> 84

<170> PERL Program

<210> 1

<211> 261

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 016233CD1

<400> 1

Met	Ala	Gly	Asp	Ser	Glu	Gln	Thr	Leu	Gln	Asn	His	Gln	Gln	Pro	1	5	10	15
Asn	Gly	Gly	Glu	Pro	Phe	Leu	Ile	Gly	Val	Ser	Gly	Gly	Thr	Ala	20	25	30	35
Ser	Gly	Lys	Ser	Ser	Val	Cys	Ala	Lys	Ile	Val	Gln	Leu	Leu	Gly	40	45	50	55
Gln	Asn	Glu	Val	Asp	Tyr	Arg	Gln	Lys	Gln	Val	Val	Ile	Leu	Ser	60	65	70	75
Gln	Asp	Ser	Phe	Tyr	Arg	Val	Leu	Thr	Ser	Glu	Gln	Lys	Ala	Lys	80	85	90	95
Ala	Leu	Lys	Gly	Gln	Phe	Asn	Phe	Asp	His	Pro	Asp	Ala	Phe	Asp	100	105	110	115
Asn	Glu	Leu	Ile	Leu	Lys	Thr	Leu	Lys	Glu	Ile	Thr	Glu	Gly	Lys	120	125	130	135
Thr	Val	Gln	Ile	Pro	Val	Tyr	Asp	Phe	Val	Ser	His	Ser	Arg	Lys	140	145	150	155
Glu	Glu	Thr	Val	Thr	Val	Tyr	Pro	Ala	Asp	Val	Val	Leu	Phe	Glu	160	165	170	175
Gly	Ile	Leu	Ala	Phe	Tyr	Ser	Gln	Glu	Val	Arg	Asp	Leu	Phe	Gln	180	185	190	195
Met	Lys	Leu	Phe	Val	Asp	Thr	Asp	Ala	Asp	Thr	Arg	Leu	Ser	Arg	200	205	210	215
Arg	Val	Leu	Arg	Asp	Ile	Ser	Glu	Arg	Gly	Arg	Asp	Leu	Glu	Gln	220	225	230	235
Ile	Leu	Ser	Gln	Tyr	Ile	Thr	Phe	Val	Lys	Pro	Ala	Phe	Glu	Glu	240	245	250	255
Phe	Cys	Leu	Pro	Thr	Lys	Lys	Tyr	Ala	Asp	Val	Ile	Ile	Pro	Arg	260	265	270	275
Gly	Ala	Asp	Asn	Leu	Val	Ala	Ile	Asn	Leu	Ile	Val	Gln	His	Ile	280	285	290	295

Gln Asp Ile Leu	215	Asn Gly Gly Pro Ser	220	Lys Arg Gln Thr Asn	225
	230		235		240
Cys Leu Asn Gly	245	Tyr Thr Pro Ser Arg	250	Lys Arg Gln Ala Ser	255
Ser Ser Ser Arg	260	Pro His			

<210> 2  
 <211> 197  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 078336CD1

<400> 2

Met Phe Tyr Leu Val	5	Val Phe Leu Asn Ser	10	Gly Asp Ile Gln Glu	15
Leu Tyr Asp Thr Thr	20	Leu Ala Leu Gly His	25	Ala Ala Ala Phe Ser	30
Asp Asp Cys Asp Leu	35	Pro Ser Ala Gln Asp	40	Ile Asn Arg Leu Val	45
Gly Leu Gln Asn Thr	50	Tyr Met Gly Tyr Leu	55	Asp Tyr Arg Lys Lys	60
Ala Ile Lys Asp Leu	65	Gly Ile Ser Pro Ser	70	Thr Cys Ser Phe Asn	75
Pro Gly Val Ile Val	80	Ala Asn Met Thr Glu	85	Trp Lys His Gln Arg	90
Ile Thr Lys Gln Leu	95	Glu Lys Trp Met Gln	100	Lys Asn Val Glu Glu	105
Asn Leu Tyr Ser Ser	110	Ser Leu Gly Gly Gly	115	Val Ala Thr Ser Pro	120
Met Leu Ile Val Phe	125	His Gly Lys Tyr Ser	130	Thr Ile Asn Pro Leu	135
Trp His Ile Arg His	140	Leu Gly Trp Asn Pro	145	Asp Ala Arg Tyr Ser	150
Glu His Phe Leu Gln	155	Glu Ala Lys Leu Leu	160	His Trp Asn Gly Arg	165
His Lys Pro Trp Asp	170	Phe Pro Ser Val His	175	Asn Asp Leu Trp Glu	180
Ser Trp Phe Val Pro	185	Asp Pro Ala Gly Ile	190	Phe Lys Leu Asn His	195
His Ser					

<210> 3  
 <211> 378  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 130117CD1

<400> 3

Met Asp Leu Ala Gly	5	Leu Leu Lys Ser Gln	10	Phe Leu Cys His Leu	15
Val Phe Cys Tyr Val	20	Phe Ile Ala Ser Gly	25	Leu Ile Ile Asn Thr	30
Ile Gln Leu Phe Thr	35	Leu Leu Leu Trp Pro	40	Ile Asn Lys Gln Leu	45
Phe Arg Lys Ile Asn	50	Cys Arg Leu Ser Tyr	55	Cys Ile Ser Ser Gln	60
Leu Val Met Leu Leu	65	Glu Trp Trp Ser Gly	70	Thr Glu Cys Thr Ile	75

Phe Thr Asp Pro Arg Ala Tyr Leu Lys Tyr Gly Lys Glu Asn Ala  
 80 85 90  
 Ile Val Val Leu Asn His Lys Phe Glu Ile Asp Phe Leu Cys Gly  
 95 100 105  
 Trp Ser Leu Ser Glu Arg Phe Gly Leu Leu Gly Gly Ser Lys Val  
 110 115 120  
 Leu Ala Lys Lys Glu Leu Ala Tyr Val Pro Ile Ile Gly Trp Met  
 125 130 135  
 Trp Tyr Phe Thr Glu Met Val Phe Cys Ser Arg Lys Trp Glu Gln  
 140 145 150  
 Asp Arg Lys Thr Val Ala Thr Ser Leu Gln His Leu Arg Asp Tyr  
 155 160 165  
 Pro Glu Lys Tyr Phe Phe Leu Ile His Cys Glu Gly Thr Arg Phe  
 170 175 180  
 Thr Glu Lys Lys His Glu Ile Ser Met Gln Val Ala Arg Ala Lys  
 185 190 195  
 Gly Leu Pro Arg Leu Lys His His Leu Leu Pro Arg Thr Lys Gly  
 200 205 210  
 Phe Ala Ile Thr Val Arg Ser Leu Arg Asn Val Val Ser Ala Val  
 215 220 225  
 Tyr Asp Cys Thr Leu Asn Phe Arg Asn Asn Glu Asn Pro Thr Leu  
 230 235 240  
 Leu Gly Val Leu Asn Gly Lys Lys Tyr His Ala Asp Leu Tyr Val  
 245 250 255  
 Arg Arg Ile Pro Leu Glu Asp Ile Pro Glu Asp Asp Asp Glu Cys  
 260 265 270  
 Ser Ala Trp Leu His Lys Leu Tyr Gln Glu Lys Asp Ala Phe Gln  
 275 280 285  
 Glu Glu Tyr Tyr Arg Thr Gly Thr Phe Pro Glu Thr Pro Met Val  
 290 295 300  
 Pro Pro Arg Arg Pro Trp Thr Leu Val Asn Trp Leu Phe Trp Ala  
 305 310 315  
 Ser Leu Val Leu Tyr Pro Phe Phe Gln Phe Leu Val Ser Met Ile  
 320 325 330  
 Arg Ser Gly Ser Ser Leu Thr Leu Ala Ser Phe Ile Leu Val Phe  
 335 340 345  
 Phe Val Ala Ser Val Gly Val Arg Trp Met Ile Gly Val Thr Glu  
 350 355 360  
 Ile Asp Lys Gly Ser Ala Tyr Gly Asn Ser Asp Ser Lys Gln Lys  
 365 370 375  
 Leu Asn Asp

<210> 4  
 <211> 285  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 267495CD1

<400> 4  
 Met Leu Tyr Leu Ile Gly Leu Gly Leu Gly Asp Ala Lys Asp Ile  
 1 5 10 15  
 Thr Val Lys Gly Leu Glu Val Val Arg Arg Cys Ser Arg Val Tyr  
 20 25 30  
 Leu Glu Ala Tyr Thr Ser Val Leu Thr Val Gly Lys Glu Ala Leu  
 35 40 45  
 Glu Glu Phe Tyr Gly Arg Lys Leu Val Val Ala Asp Arg Glu Glu  
 50 55 60  
 Val Glu Gln Glu Ala Asp Asn Ile Leu Lys Asp Ala Asp Ile Ser  
 65 70 75  
 Asp Val Ala Phe Leu Val Val Gly Asp Pro Phe Gly Ala Thr Thr  
 80 85 90  
 His Ser Asp Leu Val Leu Arg Ala Thr Lys Leu Gly Ile Pro Tyr  
 95 100 105

Arg	Val	Ile	His	Asn	Ala	Ser	Ile	Met	Asn	Ala	Val	Gly	Cys	Cys	
				110					115					120	
Gly	Leu	Gln	Leu	Tyr	Lys	Phe	Gly	Glu	Thr	Val	Ser	Ile	Val	Phe	
				125					130					135	
Trp	Thr	Asp	Thr	Trp	Arg	Pro	Glu	Ser	Phe	Phe	Asp	Lys	Val	Lys	
				140					145					150	
Lys	Asn	Arg	Gln	Asn	Gly	Met	His	Thr	Leu	Cys	Leu	Leu	Asp	Ile	
				155					160					165	
Lys	Val	Lys	Glu	Gln	Ser	Leu	Glu	Asn	Leu	Ile	Lys	Gly	Arg	Lys	
				170					175					180	
Ile	Tyr	Glu	Pro	Pro	Arg	Tyr	Met	Ser	Val	Asn	Gln	Ala	Ala	Gln	
				185					190					195	
Gln	Leu	Leu	Glu	Ile	Val	Gln	Asn	Gln	Arg	Ile	Arg	Gly	Glu	Glu	
				200					205					210	
Pro	Ala	Val	Thr	Glu	Glu	Thr	Leu	Cys	Val	Gly	Leu	Ala	Arg	Val	
				215					220					225	
Gly	Ala	Asp	Asp	Gln	Lys	Ile	Ala	Ala	Gly	Thr	Leu	Arg	Gln	Met	
				230					235					240	
Cys	Thr	Val	Asp	Leu	Gly	Glu	Pro	Leu	His	Ser	Leu	Ile	Ile	Thr	
				245					250					255	
Gly	Gly	Ser	Ile	His	Pro	Met	Glu	Met	Glu	Met	Leu	Ser	Leu	Phe	
				260					265					270	
Ser	Ile	Pro	Glu	Asn	Ser	Ser	Glu	Ser	Gln	Ser	Ile	Asn	Gly	Leu	
				275					280					285	

&lt;210&gt; 5

&lt;211&gt; 301

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 410533CD1

&lt;400&gt; 5

Met	Leu	Ala	Leu	Arg	Val	Ala	Arg	Gly	Ser	Trp	Gly	Ala	Leu	Arg	
1				5					10					15	
Gly	Ala	Ala	Trp	Ala	Pro	Gly	Thr	Arg	Pro	Ser	Lys	Arg	Arg	Ala	
				20					25					30	
Cys	Trp	Ala	Leu	Leu	Pro	Pro	Val	Pro	Cys	Cys	Leu	Gly	Cys	Leu	
				35					40					45	
Ala	Glu	Arg	Trp	Arg	Leu	Arg	Pro	Ala	Ala	Leu	Gly	Leu	Arg	Leu	
				50					55					60	
Pro	Gly	Ile	Gly	Gln	Arg	Asn	His	Cys	Ser	Gly	Ala	Gly	Lys	Ala	
				65					70					75	
Ala	Pro	Arg	Pro	Ala	Ala	Gly	Ala	Gly	Ala	Ala	Ala	Glu	Ala	Pro	
				80					85					90	
Gly	Gly	Gln	Trp	Gly	Pro	Ala	Ser	Thr	Pro	Ser	Leu	Tyr	Glu	Asn	
				95					100					105	
Pro	Trp	Thr	Ile	Pro	Asn	Met	Leu	Ser	Met	Thr	Arg	Ile	Gly	Leu	
				110					115					120	
Ala	Pro	Val	Leu	Gly	Tyr	Leu	Ile	Ile	Glu	Glu	Asp	Phe	Asn	Ile	
				125					130					135	
Ala	Leu	Gly	Val	Phe	Ala	Leu	Ala	Gly	Leu	Thr	Asp	Leu	Leu	Asp	
				140					145					150	
Gly	Phe	Ile	Ala	Arg	Asn	Trp	Ala	Asn	Gln	Arg	Ser	Ala	Leu	Gly	
				155					160					165	
Ser	Ala	Leu	Asp	Pro	Leu	Ala	Asp	Lys	Ile	Leu	Ile	Ser	Ile	Leu	
				170					175					180	
Tyr	Val	Ser	Leu	Thr	Tyr	Ala	Asp	Leu	Ile	Pro	Val	Pro	Leu	Thr	
				185					190					195	
Tyr	Met	Ile	Ile	Ser	Arg	Asp	Val	Met	Leu	Ile	Ala	Ala	Val	Phe	
				200					205					210	
Tyr	Val	Arg	Tyr	Arg	Thr	Leu	Pro	Thr	Pro	Arg	Thr	Leu	Ala	Lys	
				215					220					225	
Tyr	Phe	Asn	Pro	Cys	Tyr	Ala	Thr	Ala	Arg	Leu	Lys	Pro	Thr	Phe	

Ile Ser Lys Val	230	Thr Ala Val Gln	235	Leu Ile Leu Val Ala	240
	245		250		255
Ser Leu Ala Ala	260	Pro Val Phe Asn Tyr	265	Ala Asp Ser Ile Tyr	270
Gln Ile Leu Trp	275	Cys Phe Thr Ala Phe	280	Thr Thr Ala Ala Ser	285
Tyr Ser Tyr Tyr	290	His Tyr Gly Arg Lys	295	Thr Val Gln Val Ile	300
Asp					

<210> 6  
 <211> 253  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 852708CD1

<400> 6

Met Ala Thr Ser Gly	5	Asp Cys Pro Arg Ser	10	Glu Ser Gln Gly Glu	15
Glu Pro Ala Glu Cys	20	Ser Glu Ala Gly Leu	25	Leu Gln Glu Gly Val	30
Gln Pro Glu Glu Phe	35	Val Ala Ile Ala Asp	40	Tyr Ala Ala Thr Asp	45
Glu Thr Gln Leu Ser	50	Phe Leu Arg Gly Glu	55	Lys Ile Leu Ile Leu	60
Arg Gln Thr Thr Ala	65	Asp Trp Trp Trp Gly	70	Glu Arg Ala Gly Cys	75
Cys Gly Tyr Ile Pro	80	Ala Asn His Val Gly	85	Lys His Val Asp Glu	90
Tyr Asp Pro Glu Asp	95	Thr Trp Gln Asp Glu	100	Glu Tyr Phe Gly Ser	105
Tyr Gly Thr Leu Lys	110	Leu His Leu Glu Met	115	Leu Ala Asp Gln Pro	120
Arg Thr Thr Lys Tyr	125	His Ser Val Ile Leu	130	Gln Asn Lys Glu Ser	135
Leu Thr Asp Lys Val	140	Ile Leu Asp Val Gly	145	Cys Gly Thr Gly Ile	150
Ile Ser Leu Phe Cys	155	Ala His Tyr Ala Arg	160	Pro Arg Ala Val Tyr	165
Ala Val Glu Ala Ser	170	Glu Met Ala Gln His	175	Thr Gly Gln Leu Val	180
Leu Gln Asn Gly Phe	185	Ala Asp Ile Ile Thr	190	Val Tyr Gln Gln Lys	195
Val-Glu-Asp-Val-Val	200	Leu-Pro-Glu-Lys-Val	205	Asp-Val-Leu-Val-Ser	210
Glu Trp Met Gly Thr	215	Cys Leu Leu Val Arg	220	Ala Gly Val Arg Ala	225
Ala Gly Gly Arg Ser	230	Trp Gly Ala Ser Glu	235	His Gly Leu Gly Trp	240
Ala Asn Leu Arg Ile	245	Ser Arg Val Val Arg	250	Asp Ser Phe	

<210> 7  
 <211> 390  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 972944CD1

<400> 7

Met	Ser	Val	Arg	Val	Ala	Arg	Val	Ala	Trp	Val	Arg	Gly	Leu	Gly	15
1				5					10						15
Ala	Ser	Tyr	Arg	Arg	Gly	Ala	Ser	Ser	Phe	Pro	Val	Pro	Pro	Pro	30
				20					25						30
Gly	Ala	Gln	Gly	Val	Ala	Glu	Leu	Leu	Arg	Asp	Ala	Thr	Gly	Ala	45
				35					40						45
Glu	Glu	Glu	Ala	Pro	Trp	Ala	Ala	Thr	Glu	Arg	Arg	Met	Pro	Gly	60
				50					55						60
Gln	Cys	Ser	Val	Leu	Leu	Phe	Pro	Gly	Gln	Gly	Ser	Gln	Val	Val	75
				65					70						75
Gly	Met	Gly	Arg	Gly	Leu	Leu	Asn	Tyr	Pro	Arg	Val	Arg	Glu	Leu	90
				80					85						90
Tyr	Ala	Ala	Ala	Arg	Arg	Val	Leu	Gly	Tyr	Asp	Leu	Leu	Glu	Leu	105
				95					100						105
Ser	Leu	His	Gly	Pro	Gln	Glu	Thr	Leu	Asp	Arg	Thr	Val	His	Cys	120
				110					115						120
Gln	Pro	Ala	Ile	Phe	Val	Ala	Ser	Leu	Ala	Ala	Val	Glu	Lys	Leu	135
				125					130						135
His	His	Leu	Gln	Pro	Ser	Val	Ile	Glu	Asn	Cys	Val	Ala	Ala	Ala	150
				140					145						150
Gly	Phe	Ser	Val	Gly	Glu	Phe	Ala	Ala	Leu	Val	Phe	Ala	Gly	Ala	165
				155					160						165
Met	Glu	Phe	Ala	Glu	Gly	Leu	Tyr	Ala	Val	Lys	Ile	Arg	Ala	Glu	180
				170					175						180
Ala	Met	Gln	Glu	Ala	Ser	Glu	Ala	Val	Pro	Ser	Gly	Met	Leu	Ser	195
				185					190						195
Val	Leu	Gly	Gln	Pro	Gln	Ser	Lys	Phe	Asn	Phe	Ala	Cys	Leu	Glu	210
				200					205						210
Ala	Arg	Glu	His	Cys	Lys	Ser	Leu	Gly	Ile	Glu	Asn	Pro	Val	Cys	225
				215					220						225
Glu	Val	Ser	Asn	Tyr	Leu	Phe	Pro	Asp	Cys	Arg	Val	Ile	Ser	Gly	240
				230					235						240
His	Gln	Glu	Ala	Leu	Arg	Phe	Leu	Gln	Lys	Asn	Ser	Ser	Lys	Phe	255
				245					250						255
His	Phe	Arg	Arg	Thr	Arg	Met	Leu	Pro	Val	Ser	Gly	Ala	Phe	His	270
				260					265						270
Thr	Arg	Leu	Met	Glu	Pro	Ala	Val	Glu	Pro	Leu	Thr	Gln	Ala	Leu	285
				275					280						285
Lys	Ala	Val	Asp	Ile	Lys	Lys	Pro	Leu	Val	Ser	Val	Tyr	Ser	Asn	300
				290					295						300
Val	His	Ala	His	Arg	Tyr	Arg	His	Pro	Gly	His	Ile	His	Lys	Leu	315
				305					310						315
Leu	Ala	Gln	Gln	Leu	Val	Ser	Pro	Val	Lys	Trp	Glu	Gln	Thr	Met	330
				320					325						330
His	Ala	Ile	Tyr	Glu	Arg	Lys	Lys	Gly	Arg	Gly	Phe	Pro	Gln	Thr	345
				335					340						345
Phe	Glu	Val	Gly	Pro	Gly	Arg	Gln	Leu	Gly	Ala	Ile	Leu	Lys	Ser	360
				350					355						360
Cys	Asn	Met	Gln	Ala	Trp	Lys	Ser	Tyr	Ser	Ala	Val	Asp	Val	Leu	375
				365					370						375
Gln	Thr	Leu	Glu	His	Val	Asp	Leu	Asp	Pro	Gln	Glu	Pro	Pro	Arg	390
				380					385						390

&lt;210&gt; 8

&lt;211&gt; 373

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 997730CD1

&lt;400&gt; 8

Met Leu Ile Pro Phe Ser Met Lys Asn Cys Phe Gln Leu Leu Cys

1

5

10

15

Asn Cys Gln Val Pro Ala Ala Gly Phe Lys Lys Thr Val Lys Asn



	20		25		30									
Gly	Leu	Ile	Leu	Gln	Ser	Ile	Ser	Asn	Asp	Val	Tyr	Gln	Asn	Leu
	35		40		45									
Ala	Val	Glu	Asp	Trp	Ile	His	Asp	His	Met	Asn	Leu	Glu	Gly	Lys
	50		55		60									
Pro	Ile	Leu	Phe	Phe	Trp	Gln	Asn	Ser	Pro	Ser	Val	Val	Ile	Gly
	65		70		75									
Arg	His	Gln	Asn	Pro	Trp	Gln	Glu	Cys	Asn	Leu	Asn	Leu	Met	Arg
	80		85		90									
Glu	Glu	Gly	Ile	Lys	Leu	Ala	Arg	Arg	Arg	Ser	Gly	Gly	Gly	Thr
	95		100		105									
Val	Tyr	His	Asp	Met	Gly	Asn	Ile	Asn	Leu	Thr	Phe	Phe	Thr	Thr
	110		115		120									
Lys	Lys	Lys	Tyr	Asp	Arg	Met	Glu	Asn	Leu	Lys	Leu	Ile	Val	Arg
	125		130		135									
Ala	Leu	Asn	Ala	Val	Gln	Pro	Gln	Leu	Asp	Val	Gln	Ala	Thr	Lys
	140		145		150									
Arg	Phe	Asp	Leu	Leu	Leu	Asp	Gly	Gln	Phe	Lys	Ile	Ser	Gly	Thr
	155		160		165									
Ala	Ser	Lys	Ile	Gly	Arg	Thr	Thr	Ala	Tyr	His	His	Cys	Thr	Leu
	170		175		180									
Leu	Cys	Ser	Thr	Asp	Gly	Thr	Phe	Leu	Ser	Ser	Leu	Leu	Lys	Ser
	185		190		195									
Pro	Tyr	Gln	Gly	Ile	Arg	Ser	Asn	Ala	Thr	Ala	Ser	Ile	Pro	Ser
	200		205		210									
Leu	Val	Lys	Asn	Leu	Leu	Glu	Lys	Asp	Pro	Thr	Leu	Thr	Cys	Glu
	215		220		225									
Val	Leu	Met	Asn	Ala	Val	Ala	Thr	Glu	Tyr	Ala	Ala	Tyr	His	Gln
	230		235		240									
Ile	Asp	Asn	His	Ile	His	Leu	Ile	Asn	Pro	Thr	Asp	Glu	Thr	Leu
	245		250		255									
Phe	Pro	Gly	Ile	Asn	Ser	Lys	Ala	Lys	Glu	Leu	Gln	Thr	Trp	Glu
	260		265		270									
Trp	Ile	Tyr	Gly	Lys	Thr	Pro	Lys	Phe	Ser	Ile	Asn	Thr	Ser	Phe
	275		280		285									
His	Val	Leu	Tyr	Glu	Gln	Ser	His	Leu	Glu	Ile	Lys	Val	Phe	Ile
	290		295		300									
Asp	Ile	Lys	Asn	Gly	Arg	Ile	Glu	Ile	Cys	Asn	Ile	Glu	Ala	Pro
	305		310		315									
Asp	His	Trp	Leu	Pro	Leu	Glu	Ile	Arg	Asp	Lys	Leu	Asn	Ser	Ser
	320		325		330									
Leu	Ile	Gly	Ser	Lys	Phe	Cys	Pro	Thr	Glu	Thr	Thr	Met	Leu	Thr
	335		340		345									
Asn	Ile	Leu	Leu	Arg	Thr	Cys	Pro	Gln	Asp	His	Lys	Leu	Asn	Ser
	350		355		360									
Lys	Trp	Asn	Ile	Leu	Cys	Glu	Lys	Ile	Lys	Gly	Ile	Met		
	365		370											

&lt;210&gt; 9

&lt;211&gt; 371

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1285944CD1

&lt;400&gt; 9

Met	Ser	Phe	Arg	Lys	Val	Asn	Ile	Ile	Ile	Leu	Val	Leu	Ala	Val
1				5					10					15
Ala	Leu	Phe	Leu	Leu	Val	Leu	His	His	Asn	Phe	Leu	Ser	Leu	Ser
				20					25					30
Ser	Leu	Leu	Arg	Asn	Glu	Val	Thr	Asp	Ser	Gly	Ile	Val	Gly	Pro
				35					40					45
Gln	Pro	Ile	Asp	Phe	Val	Pro	Asn	Ala	Leu	Arg	His	Ala	Val	Asp
				50					55					60
Gly	Arg	Gln	Glu	Glu	Ile	Pro	Val	Val	Ile	Ala	Ala	Ser	Glu	Asp

65	70	75
Arg Leu Gly Gly Ala	Ile Ala Ala Ile Asn Ser Ile Gln His Asn	
80	85	90
Thr Arg Ser Asn Val	Ile Phe Tyr Ile Val Thr Leu Asn Asn Thr	
95	100	105
Ala Asp His Leu Arg	Ser Trp Leu Asn Ser Asp Ser Leu Lys Ser	
110	115	120
Ile Arg Tyr Lys Ile	Val Asn Phe Asp Pro Lys Leu Leu Glu Gly	
125	130	135
Lys Val Lys Glu Asp	Pro Asp Gln Gly Glu Ser Met Lys Pro Leu	
140	145	150
Thr Phe Ala Arg Phe	Tyr Leu Pro Ile Leu Val Pro Ser Ala Lys	
155	160	165
Lys Ala Ile Tyr Met	Asp Asp Asp Val Ile Val Gln Gly Asp Ile	
170	175	180
Leu Ala Leu Tyr Asn	Thr Ala Leu Lys Pro Gly His Ala Ala Ala	
185	190	195
Phe Ser Glu Asp Cys	Asp Ser Ala Ser Thr Lys Val Val Ile Arg	
200	205	210
Gly Ala Gly Asn Gln	Tyr Asn Tyr Ile Gly Tyr Leu Asp Tyr Lys	
215	220	225
Lys Glu Arg Ile Arg	Lys Leu Ser Met Lys Ala Ser Thr Cys Ser	
230	235	240
Phe Asn Pro Gly Val	Phe Val Ala Asn Leu Thr Glu Trp Lys Arg	
245	250	255
Gln Asn Ile Thr Asn	Gln Leu Glu Lys Trp Met Lys Leu Asn Val	
260	265	270
Glu Glu Gly Leu Tyr	Ser Arg Thr Leu Ala Gly Ser Ile Thr Thr	
275	280	285
Pro Pro Leu Leu Ile	Val Phe Tyr Gln Gln His Ser Thr Ile Asp	
290	295	300
Pro Met Trp Asn Val	Arg His Leu Gly Ser Ser Ala Gly Lys Arg	
305	310	315
Tyr Ser Pro Gln Phe	Val Lys Ala Ala Lys Leu Leu His Trp Asn	
320	325	330
Gly His Leu Lys Pro	Trp Gly Arg Thr Ala Ser Tyr Thr Asp Val	
335	340	345
Trp Glu Lys Trp Tyr	Ile Pro Asp Pro Thr Gly Lys Phe Asn Leu	
350	355	360
Ile Arg Arg Tyr Thr	Glu Ile Ser Asn Ile Lys	
365	370	

&lt;210&gt; 10

&lt;211&gt; 123

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1293207CD1

&lt;400&gt; 10

Met Phe Asn Phe Asp	Thr Phe Trp Lys Asn Phe Lys Ser Lys Leu
1	5
Gly Phe Ile Asn Trp	Asp Ala Ile Asn Lys Asn Gln Val Pro Pro
20	25
Pro Ser Thr Arg Ala	Leu Leu Tyr Phe Ser Arg Leu Trp Glu Asp
35	40
Phe Lys Gln Asn Thr	Pro Phe Leu Asn Trp Lys Ala Ile Ile Glu
50	55
Gly Ala Asp Ala Ser	Ser Leu Gln Lys Arg Ala Gly Arg Ala Asp
65	70
Gln Asn Tyr Asn Tyr	Asn Gln His Ala Tyr Pro Thr Ala Tyr Gly
80	85
Gly Lys Tyr Ser Val	Lys Thr Pro Ala Lys Gly Gly Val Ser Pro
95	100
Ser Ser Ser Ala Ser	Arg Val Gln Pro Gly Leu Leu Gln Trp Val

Lys Phe Trp 110 115 120  
 <210> 11  
 <211> 85  
 <212> PRT  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1308125CD1  
 <400> 11  
 Met Ser Ser Ser Arg Met Glu Gly Lys Ala Lys Tyr Ile Leu Pro  
 1 5 10 15  
 Thr Glu Thr Ile Tyr Val Gly Glu Met Lys Asp Gly Met Phe His  
 20 25 30  
 Gly Glu Gly Thr Leu Tyr Phe Pro Ser Gly Ser Gln Tyr Asp Ala  
 35 40 45  
 Ile Trp Glu Asn Gly Leu Ala Ile Lys Val Trp Leu Asn Ser Pro  
 50 55 60  
 Ile Trp Thr His Leu Glu Lys Ser Pro Arg Ala Ile Thr Ile Val  
 65 70 75  
 Glu Thr Ala Ser Ile Thr Gln Ser Arg Gly  
 80 85  
 <210> 12  
 <211> 184  
 <212> PRT  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1439670CD1  
 <400> 12  
 Met Lys Pro Asp Glu Thr Pro Met Phe Asp Pro Ser Leu Leu Lys  
 1 5 10 15  
 Glu Val Asp Trp Ser Gln Asn Thr Ala Thr Phe Ser Pro Ala Ile  
 20 25 30  
 Ser Pro Thr His Pro Gly Glu Gly Leu Val Leu Arg Pro Leu Cys  
 35 40 45  
 Thr Ala Asp Leu Asn Arg Gly Phe Phe Lys Val Leu Gly Gln Leu  
 50 55 60  
 Thr Glu Thr Gly Val Val Ser Pro Glu Gln Phe Met Lys Ser Phe  
 65 70 75  
 Glu His Met Lys Lys Ser Gly Asp Tyr Tyr Val Thr Val Val Glu  
 80 85 90  
 Asp Val Thr Leu Gly Gln Ile Val Ala Thr Ala Thr Leu Ile Ile  
 95 100 105  
 Glu His Lys Phe Ile His Ser Cys Ala Lys Arg Gly Arg Val Glu  
 110 115 120  
 Asp Val Val Val Ser Asp Glu Cys Arg Gly Lys Gln Leu Gly Lys  
 125 130 135  
 Leu Leu Leu Ser Thr Leu Thr Leu Leu Ser Lys Lys Leu Asn Cys  
 140 145 150  
 Tyr Lys Ile Thr Leu Glu Cys Leu Pro Gln Asn Val Gly Phe Tyr  
 155 160 165  
 Lys Lys Phe Gly Tyr Thr Val Ser Glu Glu Asn Tyr Met Cys Arg  
 170 175 180  
 Arg Phe Leu Lys  
 <210> 13  
 <211> 169  
 <212> PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1444281CD1

&lt;400&gt; 13

```

Met Ala Asn Tyr Ile His Val Pro Pro Gly Ser Pro Glu Val Pro
 1          5          10          15
Lys Leu Asn Val Thr Val Gln Asp Gln Glu Glu His Arg Cys Arg
          20          25          30
Glu Gly Ala Leu Ser Leu Leu Gln His Leu Arg Pro His Trp Asp
          35          40          45
Pro Gln Glu Val Thr Leu Gln Leu Phe Thr Asp Gly Ile Thr Asn
          50          55          60
Lys Leu Ile Gly Cys Tyr Val Gly Asn Thr Met Glu Asp Val Val
          65          70          75
Leu Val Arg Ile Tyr Gly Asn Lys Thr Glu Leu Leu Val Asp Arg
          80          85          90
Asp Glu Glu Val Lys Ser Phe Arg Val Leu Gln Ala His Gly Cys
          95          100          105
Ala Pro Gln Leu Tyr Cys Thr Phe Asn Asn Gly Leu Cys Tyr Glu
          110          115          120
Phe Ile Gln Gly Glu Ala Leu Asp Pro Lys His Val Cys Asn Pro
          125          130          135
Ala Ile Phe Ser Leu Ser Ser Leu Thr Leu Cys Lys Gly Lys Thr
          140          145          150
Thr Arg Cys Phe Gly Leu Thr Gly Cys Arg Gly Ser Arg Leu Leu
          155          160          165
Leu Ser Phe Phe

```

&lt;210&gt; 14

&lt;211&gt; 357

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1450140CD1

&lt;400&gt; 14

```

Met Gly Gly Ala Val Ser Ala Gly Glu Asp Asn Asp Asp Leu Ile
 1          5          10          15
Asp Asn Leu Lys Glu Ala Gln Tyr Ile Arg Thr Glu Arg Val Glu
          20          25          30
Gln Ala Phe Arg Ala Ile Asp Arg Gly Asp Tyr Tyr Leu Glu Gly
          35          40          45
Tyr Arg Asp Asn Ala Tyr Lys Asp Leu Ala Trp Lys His Gly Asn
          50          55          60
Ile His Leu Ser Ala Pro Cys Ile Tyr Ser Glu Val Met Glu Ala
          65          70          75
Leu Lys Leu Gln Pro Gly Leu Ser Phe Leu Asn Leu Gly Ser Gly
          80          85          90
Thr Gly Tyr Leu Ser Thr Met Val Gly Leu Ile Leu Gly Pro Phe
          95          100          105
Gly Ile Asn His Gly Ile Glu Leu His Ser Asp Val Val Glu Tyr
          110          115          120
Ala Lys Glu Lys Leu Glu Ser Phe Ile Lys Asn Ser Asp Ser Phe
          125          130          135
Asp Lys Phe Glu Phe Cys Glu Pro Ala Phe Val Val Gly Asn Cys
          140          145          150
Leu Gln Ile Ala Ser Asp Ser His Gln Tyr Asp Arg Ile Tyr Cys
          155          160          165
Gly Ala Gly Val Gln Lys Asp His Glu Asn Tyr Met Lys Ile Leu
          170          175          180
Leu Lys Val Gly Gly Ile Leu Val Met Pro Ile Glu Asp Gln Leu

```

Thr Gln Ile Met	185	Thr Gly Gln Asn	190	Thr Trp Glu Ser Lys	195
Ile Leu Ala Val	200	Val Gln Pro Ser Lys	205	Asn	210
Asp Asn Gly Lys	215	Pro Asp Ser Val Gly	220	Pro Pro Cys Ala	225
Arg Asn Leu Gln	230	Leu Ala Arg Ile	235	Thr Leu	240
Arg Asn Phe Ile	245	Asp Glu Met Gln	250	Ala Lys Gly Ile Pro	255
Arg Ala Pro Pro	260	Lys Arg Lys Arg Lys	265	Val Lys Gln Arg	270
Asn Thr Tyr Val	275	Phe Val Gly Asn Gln	280	Leu Ile Pro Gln Pro	285
Asp Ser Glu Glu	290	Asp Glu Lys Met Glu	295	Asp Ile Lys Glu	300
Glu Glu Lys Asp	305	His Asn Glu Ala Met	310	Lys Pro Glu Glu Pro	315
Gln Asn Leu Leu	320	Arg Glu Lys Ile Met	325	Lys Leu Pro Leu Pro	330
Ser Leu Lys Ala	335	Tyr Leu Thr Tyr Phe	340	Arg Asp Lys	345
	350		355		

<210> 15  
 <211> 100  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1604828CD1

Met Asn Val Arg Gly Lys Val Ile Leu Ser Met Leu Val Val Ser		
1 5 10 15		
Thr Val Ile Ile Val Phe Trp Glu Phe Ile Asn Ser Thr Glu Asp	20	25
Ser Phe Leu Trp Ile Tyr His Ser Lys Asn Pro Glu Val Asp Asp	35	40
Ser Ser Ala Gln Lys Gly Trp Trp Phe Leu Ser Trp Phe Asn Asn	50	55
Gly Ile His Asn Tyr Gln Gln Gly Glu Glu Asp Ile Asp Lys Glu	65	70
Lys Gly Arg Glu Glu Thr Lys Gly Arg Lys Met Thr Gln Gln Ser	80	85
Phe Gly Tyr Gly Thr Gly Leu Ile Gln Thr	95	100

<210> 16  
 <211> 199  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1644023CD1

Met Lys Thr Phe Ile Ile Gly Ile Ser Gly Val Thr Asn Ser Gly		
1 5 10 15		
Lys Thr Thr Leu Ala Lys Asn Leu Gln Lys His Leu Pro Asn Cys	20	25
Ser Val Ile Ser Gln Asp Asp Phe Phe Lys Pro Glu Ser Glu Ile	35	40
Glu Thr Asp Lys Asn Gly Phe Leu Gln Tyr Asp Val Leu Glu Ala	50	55

Leu Asn Met Glu Lys Met Met Ser Ala Ile Ser Cys Trp Met Glu  
 65 70 75  
 Ser Ala Arg His Ser Val Val Ser Thr Asp Gln Glu Ser Ala Glu  
 80 85 90  
 Glu Ile Pro Ile Leu Ile Ile Glu Gly Phe Leu Leu Phe Asn Tyr  
 95 100 105  
 Lys Pro Leu Asp Thr Ile Trp Asn Arg Ser Tyr Phe Leu Thr Ile  
 110 115 120  
 Pro Tyr Glu Glu Cys Lys Arg Arg Arg Ser Thr Arg Val Tyr Gln  
 125 130 135  
 Pro Pro Asp Ser Pro Gly Tyr Phe Asp Gly His Val Trp Pro Met  
 140 145 150  
 Tyr Leu Lys Tyr Arg Gln Glu Met Gln Asp Ile Thr Trp Glu Val  
 155 160 165  
 Val Tyr Leu Asp Gly Thr Lys Ser Glu Glu Asp Leu Phe Leu Gln  
 170 175 180  
 Val Tyr Glu Asp Leu Ile Gln Glu Leu Ala Lys Gln Lys Cys Leu  
 185 190 195  
 Gln Val Thr Ala

<210> 17  
 <211> 244  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1723402CD1

<400> 17  
 Met Glu Leu Thr Ile Phe Ile Leu Arg Leu Ala Ile Tyr Ile Leu  
 1 5 10 15  
 Thr Phe Pro Leu Tyr Leu Leu Asn Phe Leu Gly Leu Trp Ser Trp  
 20 25 30  
 Ile Cys Lys Lys Trp Phe Pro Tyr Phe Leu Val Arg Phe Thr Val  
 35 40 45  
 Ile Tyr Asn Glu Gln Met Ala Ser Lys Lys Arg Glu Leu Phe Ser  
 50 55 60  
 Asn Leu Gln Glu Phe Ala Gly Pro Ser Gly Lys Leu Ser Leu Leu  
 65 70 75  
 Glu Val Gly Cys Gly Thr Gly Ala Asn Phe Lys Phe Tyr Pro Pro  
 80 85 90  
 Gly Cys Arg Val Thr Cys Ile Asp Pro Asn Pro Asn Phe Glu Lys  
 95 100 105  
 Phe Leu Ile Lys Ser Ile Ala Glu Asn Arg His Leu Gln Phe Glu  
 110 115 120  
 Arg Phe Val Val Ala Ala Gly Glu Asn Met His Gln Val Ala Asp  
 125 130 135  
 Gly Ser Val Asp Val Val Val Cys Thr Leu Val Leu Cys Ser Val  
 140 145 150  
 Lys Asn Gln Glu Arg Ile Leu Arg Glu Val Cys Arg Val Leu Arg  
 155 160 165  
 Pro Gly Gly Ala Phe Tyr Phe Met Glu His Val Ala Ala Glu Cys  
 170 175 180  
 Ser Thr Trp Asn Tyr Phe Trp Gln Gln Val Leu Asp Pro Ala Trp  
 185 190 195  
 His Leu Leu Phe Asp Gly Cys Asn Leu Thr Arg Glu Ser Trp Lys  
 200 205 210  
 Ala Leu Glu Arg Ala Ser Phe Ser Lys Leu Lys Leu Gln His Ile  
 215 220 225  
 Gln Ala Pro Leu Ser Trp Glu Leu Val Arg Pro His Ile Tyr Gly  
 230 235 240  
 Tyr Ala Val Lys

<210> 18

<211> 358  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1740585CD1

<400> 18  
 Met Lys Thr Ala Glu Asn Ile Arg Gly Thr Gly Ser Asp Gly Pro  
 1 5 10 15  
 Arg Lys Arg Gly Leu Cys Val Leu Cys Gly Leu Pro Ala Ala Gly  
 20 25 30  
 Lys Ser Thr Phe Ala Arg Ala Leu Ala His Arg Leu Gln Gln Glu  
 35 40 45  
 Gln Gly Trp Ala Ile Gly Val Val Ala Tyr Asp Asp Val Met Pro  
 50 55 60  
 Asp Ala Phe Leu Ala Gly Ala Arg Ala Arg Pro Ala Pro Ser Gln  
 65 70 75  
 Trp Lys Leu Leu Arg Gln Glu Leu Leu Lys Tyr Leu Glu Tyr Phe  
 80 85 90  
 Leu Met Ala Val Ile Asn Gly Cys Gln Met Ser Val Pro Pro Asn  
 95 100 105  
 Arg Thr Glu Ala Met Trp Glu Asp Phe Ile Thr Cys Leu Lys Asp  
 110 115 120  
 Gln Asp Leu Ile Phe Ser Ala Ala Phe Glu Ala Gln Ser Cys Tyr  
 125 130 135  
 Leu Leu Thr Lys Thr Ala Val Ser Arg Pro Leu Phe Leu Val Leu  
 140 145 150  
 Asp Asp Asn Phe Tyr Tyr Gln Ser Met Arg Tyr Glu Val Tyr Gln  
 155 160 165  
 Leu Ala Arg Lys Tyr Ser Leu Gly Phe Cys Gln Leu Phe Leu Asp  
 170 175 180  
 Cys Pro Leu Glu Thr Cys Leu Gln Arg Asn Gly Gln Arg Pro Gln  
 185 190 195  
 Ala Leu Pro Pro Glu Thr Ile His Leu Met Arg Arg Lys Leu Glu  
 200 205 210  
 Lys Pro Asn Pro Glu Lys Asn Ala Trp Glu His Asn Ser Leu Thr  
 215 220 225  
 Ile Pro Ser Pro Ala Cys Ala Ser Glu Ala Ser Leu Glu Val Thr  
 230 235 240  
 Asp Leu Leu Leu Thr Ala Leu Glu Asn Pro Val Lys Tyr Ala Glu  
 245 250 255  
 Asp Asn Met Glu Gln Lys Asp Thr Asp Arg Ile Ile Cys Ser Thr  
 260 265 270  
 Asn Ile Leu His Lys Thr Asp Gln Thr Leu Arg Arg Ile Val Ser  
 275 280 285  
 Gln Thr Met Lys Glu Ala Lys Asp Glu Gln Val Leu Pro His Asn  
 290 295 300  
 Leu Lys Leu Leu Ala Glu Glu Leu Asn Lys Leu Lys Ala Glu Phe  
 305 310 315  
 Leu Glu Asp Leu Lys Gln Gly Asn Lys Lys Tyr Leu Cys Phe Gln  
 320 325 330  
 Gln Thr Ile Asp Ile Pro Asp Val Ile Ser Phe Phe His Tyr Glu  
 335 340 345  
 Lys Asp Asn Ile Val Gln Lys Tyr Phe Ser Lys Gln His  
 350 355

<210> 19  
 <211> 302  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1810925CD1

&lt;400&gt; 19

```

Met Ile Leu Leu Asn Asn Ser His Lys Leu Leu Ala Leu Tyr Lys
 1          5          10          15
Ser Leu Ala Arg Ser Ile Pro Glu Ser Leu Lys Val Tyr Gly Ser
          20          25          30
Val Tyr His Ile Asn His Gly Asn Pro Phe Asn Met Glu Val Leu
          35          40          45
Val Asp Ser Trp Pro Glu Tyr Gln Met Val Ile Ile Arg Pro Gln
          50          55          60
Lys Gln Glu Met Thr Asp Asp Met Asp Ser Tyr Thr Asn Val Tyr
          65          70          75
Arg Met Phe Ser Lys Glu Pro Gln Lys Ser Glu Glu Val Leu Lys
          80          85          90
Asn Cys Glu Ile Val Asn Trp Lys Gln Arg Leu Gln Ile Gln Gly
          95          100          105
Leu Gln Glu Ser Leu Gly Glu Gly Ile Arg Val Ala Thr Phe Ser
          110          115          120
Lys Ser Val Lys Val Glu His Ser Arg Ala Leu Leu Leu Val Thr
          125          130          135
Glu Asp Ile Leu Lys Leu Asn Ala Ser Ser Lys Ser Lys Leu Gly
          140          145          150
Ser Trp Ala Glu Thr Gly His Pro Asp Asp Glu Phe Glu Ser Glu
          155          160          165
Thr Pro Asn Phe Lys Tyr Ala Gln Leu Asp Val Ser Tyr Ser Gly
          170          175          180
Leu Val Asn Asp Asn Trp Lys Arg Gly Lys Asn Glu Arg Ser Leu
          185          190          195
His Tyr Ile Lys Arg Cys Ile Glu Asp Leu Pro Ala Ala Cys Met
          200          205          210
Leu Gly Pro Glu Gly Val Pro Val Ser Trp Val Thr Met Asp Pro
          215          220          225
Ser Cys Glu Val Gly Met Ala Tyr Ser Met Glu Lys Tyr Arg Arg
          230          235          240
Thr Gly Asn Met Ala Arg Val Met Val Arg Tyr Met Lys Tyr Leu
          245          250          255
Arg Gln Lys Asn Ile Pro Phe Tyr Ile Ser Val Leu Glu Glu Asn
          260          265          270
Glu Asp Ser Arg Arg Phe Val Gly Gln Phe Gly Phe Phe Glu Ala
          275          280          285
Ser Cys Glu Trp His Gln Trp Thr Cys Tyr Pro Gln Asn Leu Val
          290          295          300
Pro Phe

```

&lt;210&gt; 20

&lt;211&gt; 234

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1915064CD1

&lt;400&gt; 20

```

Met Ser Ser Glu Val Ser Ala Arg Arg Asp Ala Lys Lys Leu Val
 1          5          10          15
Arg Ser Pro Ser Gly Leu Arg Met Val Pro Glu His Arg Ala Phe
          20          25          30
Gly Ser Pro Phe Gly Leu Glu Glu Pro Gln Trp Val Pro Asp Lys
          35          40          45
Glu Cys Arg Arg Cys Met Gln Cys Asp Ala Lys Phe Asp Phe Leu
          50          55          60
Thr Arg Lys His His Cys Arg Arg Cys Gly Lys Cys Phe Cys Asp
          65          70          75
Arg Cys Cys Ser Gln Lys Val Pro Leu Arg Arg Met Cys Phe Val
          80          85          90
Asp Pro Val Arg Gln Cys Ala Glu Cys Ala Leu Val Ser Leu Lys

```



	95		100		105
Glu Ala Glu Phe	Tyr Asp Lys Gln Leu	Lys Val Leu Leu Ser	Gly		
	110		115		120
Ala Thr Phe Leu	Val Thr Phe Gly Asn	Ser Glu Lys Pro Glu	Thr		
	125		130		135
Met Thr Cys Arg	Leu Ser Asn Asn Gln	Arg Tyr Leu Phe Leu	Asp		
	140		145		150
Gly Asp Ser His	Tyr Glu Ile Glu Ile	Val His Ile Ser Thr	Val		
	155		160		165
Gln Ile Leu Thr	Glu Gly Phe Pro Pro	Gly Gly Gly Asn Ala	Arg		
	170		175		180
Ala Thr Gly Met	Phe Leu Gln Tyr Thr	Val Pro Gly Thr Glu	Gly		
	185		190		195
Val Thr Gln Leu	Lys Leu Thr Val Val	Glu Asp Val Thr Val	Gly		
	200		205		210
Arg Arg Gln Ala	Val Ala Trp Leu Val	Ala Met His Lys Ala	Ala		
	215		220		225
Lys Leu Leu Tyr	Glu Ser Arg Asp Gln				
	230				

&lt;210&gt; 21

&lt;211&gt; 403

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2185608CD1

&lt;400&gt; 21

Met Ala Gly Ala	Ala Thr Gln Ala Ser	Leu Glu Ser Ala Pro	Arg
1	5	10	15
Ile Met Arg Leu	Val Ala Glu Cys Ser	Arg Ser Arg Ala Arg	Ala
	20	25	30
Gly Glu Leu Trp	Leu Pro His Gly Thr	Val Ala Thr Pro Val	Phe
	35	40	45
Met Pro Val Gly	Thr Gln Ala Thr Met	Lys Gly Ile Thr Thr	Glu
	50	55	60
Gln Leu Asp Ala	Leu Gly Cys Arg Ile	Cys Leu Gly Asn Thr	Tyr
	65	70	75
His Leu Gly Leu	Arg Pro Gly Pro Glu	Leu Ile Gln Lys Ala	Asn
	80	85	90
Gly Leu His Gly	Phe Met Asn Trp Pro	His Asn Leu Leu Thr	Asp
	95	100	105
Ser Gly Gly Phe	Gln Met Val Ser Leu	Val Ser Leu Ser Glu	Val
	110	115	120
Thr Glu Glu Gly	Val Arg Phe Arg Ser	Pro Tyr Asp Gly Asn	Glu
	125	130	135
Thr Leu Leu Ser	Pro Glu Lys Ser Val	Gln Ile Gln Asn Ala	Leu
	140	145	150
Gly Ser Asp Ile	Ile Met Gln Leu Asp	Asp Val Val Ser Ser	Thr
	155	160	165
Val Thr Gly Pro	Arg Val Glu Glu Ala	Met Tyr Arg Ser Ile	Arg
	170	175	180
Trp Leu Asp Arg	Cys Ile Ala Ala His	Gln Arg Pro Asp Lys	Gln
	185	190	195
Asn Leu Phe Ala	Ile Ile Gln Gly Gly	Leu Asp Ala Asp Leu	Arg
	200	205	210
Ala Thr Cys Leu	Glu Glu Met Thr Lys	Arg Asp Val Pro Gly	Phe
	215	220	225
Ala Ile Gly Gly	Leu Ser Gly Gly Glu	Ser Lys Ser Gln Phe	Trp
	230	235	240
Arg Met Val Ala	Leu Ser Thr Ser Arg	Leu Pro Lys Asp Lys	Pro
	245	250	255
Arg Tyr Leu Met	Gly Val Gly Tyr Ala	Thr Asp Leu Val Val	Cys
	260	265	270
Val Ala Leu Gly	Cys Asp Met Phe Asp	Cys Val Phe Pro Thr	Arg

Thr Ala Arg Phe	275	Ser Ala Leu Val	280	Thr Gly Asn Leu	285
Leu Arg Lys Lys	290	Val Phe Glu Lys Asp	295	Gly Pro Ile Asp	300
Glu Cys Thr Cys	305	Pro Thr Cys Gln Lys	310	Ser Arg Ala Phe	315
His Ala Leu Leu	320	His Ser Asp Asn Thr	325	Ala Leu His His	330
Thr Val His Asn	335	Ile Ala Tyr Gln Leu	340	Leu Met Ser Ala	345
Arg Thr Ser Ile	350	Val Glu Lys Arg Phe	355	Asp Phe Val Arg	360
Phe Met Gly Ala	365	Met Tyr Gly Asp Pro	370	Thr Leu Cys Pro	375
Ala Thr Asp Ala	380	Leu Ala Ser Val Gly	385	Thr Leu Gly	390
	395		400		

&lt;210&gt; 22

&lt;211&gt; 487

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2228862CD1

&lt;400&gt; 22

Met Arg Arg Gly	1	Glu Arg Arg Asp	5	Ala Gly Arg Pro	10	Arg Pro Arg Pro	15
Ser Pro Val Pro	20	Ala Gly Arg Ala	25	Ser Leu Glu Glu	30	Pro Pro Asp	35
Gly Pro Ser Ala	35	Gly Gln Ala Thr	40	Gly Pro Gly Glu	45	Gly Arg Arg	50
Ser Thr Glu Ser	50	Glu Val Tyr Asp	55	Asp Gly Thr Asn	60	Thr Phe Phe	65
Trp Arg Ala His	65	Thr Leu Thr Val	70	Leu Phe Ile Leu	75	Thr Cys Thr	80
Leu Gly Tyr Val	80	Thr Leu Leu Glu	85	Glu Thr Pro Gln	90	Asp Thr Ala	95
Tyr Asn Thr Lys	95	Arg Gly Ile Val	100	Ala Ser Ile Leu	105	Val Phe Leu	110
Cys Phe Gly Val	110	Thr Gln Ala Lys	115	Asp Gly Pro Phe	120	Ser Arg Pro	125
His Pro Ala Tyr	125	Trp Arg Phe Trp	130	Leu Cys Val Ser	135	Val Val Tyr	140
Glu Leu Phe Leu	140	Ile Phe Ile Leu	145	Phe Gln Thr Val	150	Gln Asp Gly	155
Arg Gln Phe Leu	155	Lys Tyr Val Asp	160	Pro Lys Leu Gly	165	Val Pro Leu	170
Pro Glu Arg Asp	170	Tyr Gly Gly Asn	175	Cys Leu Ile Tyr	180	Asp Pro Asp	185
Asn Glu Thr Asp	185	Pro Phe His Asn	190	Ile Trp Asp Lys	195	Leu Asp Gly	200
Phe Val Pro Ala	200	His Phe Leu Gly	205	Trp Tyr Leu Lys	210	Thr Leu Met	215
Ile Arg Asp Trp	215	Trp Met Cys Met	220	Ile Ile Ser Val	225	Met Phe Glu	230
Phe Leu Glu Tyr	230	Ser Leu Glu His	235	Gln Leu Pro Asn	240	Phe Ser Glu	245
Cys Trp Trp Asp	245	His Trp Ile Met	250	Asp Val Leu Val	255	Cys Asn Gly	260
Leu Gly Ile Tyr	260	Cys Gly Met Lys	265	Thr Leu Glu Trp	270	Leu Ser Leu	275
Lys Thr Tyr Lys	275	Trp Gln Gly Leu	280	Trp Asn Ile Pro	285	Thr Tyr Lys	290
Gly Lys Met Lys		Arg Ile Ala Phe		Gln Phe Thr Pro		Tyr Ser Trp	

Val Arg Phe Glu	290	Trp Lys Pro Ala Ser	295	Ser Leu Arg Arg Trp	300
Ala Val Cys Gly	305	Ile Ile Leu Val Phe	310	Leu Leu Ala Glu Leu	315
Thr Phe Tyr Leu	320	Lys Phe Val Leu Trp	325	Met Pro Pro Glu His	330
Leu Val Leu Leu	335	Arg Leu Val Phe Phe	340	Val Asn Val Gly Gly	345
Ala Met Arg Glu	350	Ile Tyr Asp Phe Met	355	Asp Asp Pro Lys Pro	360
Lys Lys Leu Gly	365	Pro Gln Ala Trp Leu	370	Val Ala Ala Ile Thr	375
Thr Glu Leu Leu	380	Ile Val Val Lys Tyr	385	Asp Pro His Thr Leu	390
Leu Ser Leu Pro	395	Phe Tyr Ile Ser Gln	400	Cys Trp Thr Leu Gly	405
Val Leu Ala Leu	410	Thr Trp Thr Val Trp	415	Arg Phe Phe Leu Arg	420
Ile Thr Leu Arg	425	Tyr Lys Glu Thr Arg	430	Trp Gln Lys Trp Gln	435
Lys Asp Asp Gln	440	Gly Ser Thr Val Gly	445	Asn Gly Asp Gln His	450
Leu Gly Leu Asp	455	Glu Asp Leu Leu Gly	460	Pro Gly Val Ala Glu	465
Glu Gly Ala Pro	470	Thr Pro Asn	475		480
	485				

&lt;210&gt; 23

&lt;211&gt; 246

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2235577CD1

&lt;400&gt; 23

Met Asn Asp Met	Met	Ser Leu Gly Ile	His	Arg Val Trp Lys	Asp
1	5		10		15
Leu Leu Leu Trp	Lys	Met His Pro Leu	Pro	Gly Thr Gln Leu	Leu
	20		25		30
Asp Val Ala Gly	Gly	Thr Gly Asp Ile	Ala	Phe Arg Phe Leu	Asn
	35		40		45
Tyr Val Gln Ser	Gln	His Gln Arg Lys	Gln	Lys Arg Gln Leu	Arg
	50		55		60
Ala Gln Gln Asn	Leu	Ser Trp Glu Glu	Ile	Ala Lys Glu Tyr	Gln
	65		70		75
Asn Glu Glu Asp	Ser	Leu Gly Gly Ser	Arg	Val Val Val Cys	Asp
	80		85		90
Ile Asn Lys Glu	Met	Leu Lys Val Gly	Lys	Gln Lys Ala Leu	Ala
	95		100		105
Gln Gly Tyr Arg	Ala	Gly Leu Ala Trp	Val	Leu Gly Asp Ala	Glu
	110		115		120
Glu Leu Pro Phe	Asp	Asp Asp Lys Phe	Asp	Ile Tyr Thr Ile	Ala
	125		130		135
Phe Gly Ile Arg	Asn	Val Thr His Ile	Asp	Gln Ala Leu Gln	Glu
	140		145		150
Ala His Arg Val	Leu	Lys Pro Gly Gly	Arg	Phe Leu Cys Leu	Glu
	155		160		165
Phe Ser Gln Val	Asn	Asn Pro Leu Ile	Ser	Arg Leu Tyr Asp	Leu
	170		175		180
Tyr Ser Phe Gln	Val	Ile Pro Val Leu	Gly	Glu Val Ile Ala	Gly
	185		190		195
Asp Trp Lys Ser	Tyr	Gln Tyr Leu Val	Glu	Ser Ile Arg Arg	Phe
	200		205		210
Pro Ser Gln Glu	Glu	Phe Lys Asp Met	Ile	Glu Asp Ala Gly	Phe

His Lys Val Thr	215	220	225
Tyr Glu Ser Leu Thr	230	Ser Gly Ile Val Ala	Ile
His Ser Gly Phe Lys Leu	245	235	240

<210> 24  
 <211> 410  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2271680CD1

<400> 24

Met Trp Ser Gly Arg	Lys Leu Gly Ser	Ser Gly Gly Trp Phe Leu
1 5	10	15
Arg Val Leu Gly Pro	Gly Gly Cys Asn Thr	Lys Ala Ala Arg Pro
20	25	30
Leu Ile Ser Ser Ala	Val Tyr Val Lys Asn	Gln Leu Ser Gly Thr
35	40	45
Leu Gln Ile Lys Pro	Gly Val Phe Asn Glu	Tyr Arg Thr Ile Trp
50	55	60
Phe Lys Ser Tyr Arg	Thr Ile Phe Ser Cys	Leu Asn Arg Ile Lys
65	70	75
Ser Phe Arg Trp Ser	Phe Thr Ser Val Ala	Gln Ala Gly Val Gln
80	85	90
Trp Cys Asp Leu Gly	Ser Leu Gln Pro Pro	Pro Gly Phe Lys
95	100	105
Arg Phe Ser Cys Leu	Ser Leu Leu Ser His	Trp Asp Tyr Arg Tyr
110	115	120
Pro Trp Ala Arg Leu	Tyr Ser Thr Ser Gln	Thr Thr Val Asp Ser
125	130	135
Gly Glu Val Lys Thr	Phe Leu Ala Leu Ala	His Lys Trp Trp Asp
140	145	150
Glu Gln Gly Val Tyr	Ala Pro Leu His Ser	Met Asn Asp Leu Arg
155	160	165
Val Pro Phe Ile Arg	Asp Asn Leu Leu Lys	Thr Ile Pro Asn His
170	175	180
Gln Pro Gly Lys Pro	Leu Leu Gly Met Lys	Ile Leu Asp Val Gly
185	190	195
Cys Gly Gly Gly Leu	Leu Thr Glu Pro Leu	Gly Arg Leu Gly Ala
200	205	210
Ser Val Ile Gly Ile	Asp Pro Val Asp Glu	Asn Ile Lys Thr Ala
215	220	225
Gln Cys His Lys Ser	Phe Asp Pro Val Leu	Asp Lys Arg Ile Glu
230	235	240
Tyr Arg Val Cys Ser	Leu Glu Glu Ile Val	Glu Glu Thr Ala Glu
245	250	255
Thr Phe Asp Ala Val	Val Ala Ser Glu Val	Val Glu His Val Ile
260	265	270
Asp Leu Glu Thr Phe	Leu Gln Cys Cys Cys	Gln Val Leu Lys Pro
275	280	285
Gly Gly Ser Leu Phe	Ile Thr Thr Ile Asn	Lys Thr Gln Leu Ser
290	295	300
Tyr Ala Leu Gly Ile	Val Phe Ser Glu Gln	Ile Ala Gly Ile Val
305	310	315
Pro Lys Gly Thr His	Thr Trp Glu Lys Phe	Val Ser Pro Glu Thr
320	325	330
Leu Glu Ser Ile Leu	Glu Ser Asn Gly Leu	Ser Val Gln Thr Val
335	340	345
Val Gly Met Leu Tyr	Asn Pro Phe Ser Gly	Tyr Trp His Trp Ser
350	355	360
Glu Asn Thr Ser Leu	Asn Tyr Ala Ala His	Ala Val Lys Ser Arg
365	370	375
Val Gln Glu His Pro	Ala Ser Ala Glu Phe	Val Leu Lys Gly Glu

Thr Glu Glu Leu	380	Gln Ala Asn Ala Cys	385	Thr Asn Pro Ala Val	390
	395		400		405
Glu Lys Leu Lys	Lys				
	410				

<210> 25  
 <211> 253  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2325603CD1

<400> 25  
 Met Asn Phe Ser Gly Gly Gly Arg Gln Glu Ala Ala Gly Ser Arg  
 1 5 10 15  
 Ser Arg Arg Ala Pro Arg Pro Arg Glu Gln Asp Arg Asp Val Gln  
 20 25 30  
 Leu Ser Lys Ala Leu Ser Tyr Ala Leu Arg His Gly Ala Leu Lys  
 35 40 45  
 Leu Gly Leu Pro Met Gly Ala Asp Gly Phe Val Pro Leu Gly Thr  
 50 55 60  
 Leu Leu Gln Leu Pro Gln Phe Arg Gly Phe Ser Ala Glu Asp Val  
 65 70 75  
 Gln Arg Val Val Asp Thr Asn Arg Lys Gln Arg Phe Ala Leu Gln  
 80 85 90  
 Leu Gly Asp Pro Ser Thr Gly Leu Leu Ile Arg Ala Asn Gln Gly  
 95 100 105  
 His Ser Leu Gln Val Pro Lys Leu Glu Leu Met Pro Leu Glu Thr  
 110 115 120  
 Pro Gln Ala Leu Pro Pro Met Leu Val His Gly Thr Phe Trp Lys  
 125 130 135  
 His Trp Pro Ser Ile Leu Leu Lys Gly Leu Ser Cys Gln Gly Arg  
 140 145 150  
 Thr His Ile His Leu Ala Pro Gly Leu Pro Gly Asp Pro Gly Ile  
 155 160 165  
 Ile Ser Gly Met Arg Ser His Cys Glu Ile Ala Val Phe Ile Asp  
 170 175 180  
 Gly Pro Leu Ala Leu Ala Asp Gly Ile Pro Phe Phe Arg Ser Ala  
 185 190 195  
 Asn Gly Val Ile Leu Thr Pro Gly Asn Thr Asp Gly Phe Leu Leu  
 200 205 210  
 Pro Lys Tyr Phe Lys Glu Ala Leu Gln Leu Arg Pro Thr Arg Lys  
 215 220 225  
 Pro Leu Ser Leu Ala Gly Asp Glu Glu Thr Glu Cys Gln Ser Ser  
 230 235 240  
 Pro Lys His Ser Ser Arg Glu Arg Arg Arg Ile Gln Gln  
 245 250

<210> 26  
 <211> 303  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2356055CD1

<400> 26  
 Met Lys Gly Gly Phe Thr Gly Gly Asp Glu Tyr Gln Lys His Phe  
 1 5 10 15  
 Leu Pro Arg Asp Tyr Leu Ala Thr Tyr Tyr Ser Phe Asn Gly Ser  
 20 25 30  
 Pro Ser Pro Glu Ala Glu Met Leu Lys Phe Asn Leu Glu Cys Leu  
 35 40 45

```

His Lys Thr Phe Gly Pro Gly Gly Leu Gln Gly Asp Thr Leu Ile
50 55 60
Asp Ile Gly Ser Gly Pro Thr Ile Tyr Gln Val Leu Ala Ala Cys
65 70 75
Asp Ser Phe Gln Asp Ile Thr Leu Ser Asp Phe Thr Asp Arg Asn
80 85 90
Arg Glu Glu Leu Glu Lys Trp Leu Lys Lys Glu Pro Gly Ala Tyr
95 100 105
Asp Trp Thr Pro Ala Val Lys Phe Ala Cys Glu Leu Glu Gly Asn
110 115 120
Ser Gly Arg Trp Glu Glu Lys Glu Glu Lys Leu Arg Ala Ala Val
125 130 135
Lys Arg Val Leu Lys Cys Asp Val His Leu Gly Asn Pro Leu Ala
140 145 150
Pro Ala Val Leu Pro Leu Ala Asp Cys Val Leu Thr Leu Leu Ala
155 160 165
Met Glu Cys Ala Cys Cys Ser Leu Asp Ala Tyr Arg Ala Ala Leu
170 175 180
Cys Asn Leu Ala Ser Leu Leu Lys Pro Gly Gly His Leu Val Thr
185 190 195
Thr Val Thr Leu Arg Leu Pro Ser Tyr Val Val Gly Lys Arg Glu
200 205 210
Phe Ser Cys Val Ala Leu Glu Lys Glu Glu Val Ala Ala Arg Gln
215 220 225
Cys Pro Gly Glu Glu Ile Ala Lys Glu Arg Arg Leu Gln Met Pro
230 235 240
Pro Pro Cys Asp Val Arg Thr Ser Leu Ser Glu Arg Ser Gly Gln
245 250 255
Asp Thr Gly Lys Arg His Arg Ile Gln Thr Arg Gly Ser Ala Pro
260 265 270
Trp Thr Ala Gln Cys Arg Glu Ser Ala Gly Cys Leu Glu Gly Glu
275 280 285
Ser Arg Gln Gly Cys Glu Gly Ile Phe Gly Cys Cys Gly Ser Cys
290 295 300
Ser Thr Leu

```

&lt;210&gt; 27

&lt;211&gt; 307

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2448909CD1

&lt;400&gt; 27

```

Met Gln Lys Gly Lys Gly Arg Thr Ser Arg Ile Arg Arg Arg Lys
1 5 10 15
Leu Cys Gly Ser Ser Glu Ser Arg Gly Val Asn Glu Ser His Lys
20 25 30
Ser Glu Phe Ile Glu Leu Arg Lys Trp Leu Lys Ala Arg Lys Phe
35 40 45
Gln Asp Ser Asn Leu Ala Pro Ala Cys Phe Pro Gly Thr Gly Arg
50 55 60
Gly Leu Met Ser Gln Thr Ser Leu Gln Glu Gly Gln Met Ile Ile
65 70 75
Ser Leu Pro Glu Ser Cys Leu Leu Thr Thr Asp Thr Val Ile Arg
80 85 90
Ser Tyr Leu Gly Ala Tyr Ile Thr Lys Trp Lys Pro Pro Pro Ser
95 100 105
Pro Leu Leu Ala Leu Cys Thr Phe Leu Val Ser Glu Lys His Ala
110 115 120
Gly His Arg Ser Leu Trp Lys Pro Tyr Leu Glu Ile Leu Pro Lys
125 130 135
Ala Tyr Thr Cys Pro Val Cys Leu Glu Pro Glu Val Val Asn Leu
140 145 150

```

Leu	Pro	Lys	Ser	Leu	Lys	Ala	Lys	Ala	Glu	Glu	Gln	Arg	Ala	His
				155					160					165
Val	Gln	Glu	Phe	Phe	Ala	Ser	Ser	Arg	Asp	Phe	Phe	Ser	Ser	Leu
				170					175					180
Gln	Pro	Leu	Phe	Ala	Glu	Ala	Val	Asp	Ser	Ile	Phe	Ser	Tyr	Ser
				185					190					195
Ala	Leu	Leu	Trp	Ala	Trp	Cys	Thr	Val	Asn	Thr	Arg	Ala	Val	Tyr
				200					205					210
Leu	Arg	Pro	Arg	Gln	Arg	Glu	Cys	Leu	Ser	Ala	Glu	Pro	Asp	Thr
				215					220					225
Cys	Ala	Leu	Ala	Pro	Tyr	Leu	Asp	Leu	Leu	Asn	His	Ser	Pro	His
				230					235					240
Val	Gln	Val	Lys	Ala	Ala	Phe	Asn	Glu	Glu	Thr	His	Ser	Tyr	Glu
				245					250					255
Ile	Arg	Thr	Thr	Ser	Arg	Trp	Arg	Lys	His	Glu	Glu	Val	Phe	Ile
				260					265					270
Cys	Tyr	Gly	Pro	His	Asp	Asn	Gln	Arg	Leu	Phe	Leu	Glu	Tyr	Gly
				275					280					285
Phe	Val	Ser	Val	His	Asn	Pro	His	Ala	Cys	Val	Tyr	Val	Ser	Arg
				290					295					300
Gly	Trp	Asn	Gln	Leu	Cys	Ser								305

<210> 28  
 <211> 169  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2631212CD1

<400>	28													
Met	Lys	Gly	Ser	Arg	Ile	Glu	Leu	Gly	Asp	Val	Thr	Pro	His	Asn
1				5					10					15
Ile	Lys	Gln	Leu	Lys	Arg	Leu	Asn	Gln	Val	Ile	Phe	Pro	Val	Ser
				20					25					30
Tyr	Asn	Asp	Lys	Phe	Tyr	Lys	Asp	Val	Leu	Glu	Val	Gly	Glu	Leu
				35					40					45
Ala	Lys	Leu	Ala	Tyr	Phe	Asn	Asp	Ile	Ala	Val	Gly	Ala	Val	Cys
				50					55					60
Cys	Arg	Val	Asp	His	Ser	Gln	Asn	Gln	Lys	Arg	Leu	Tyr	Ile	Met
				65					70					75
Thr	Leu	Gly	Cys	Leu	Ala	Pro	Tyr	Arg	Arg	Leu	Gly	Ile	Gly	Thr
				80					85					90
Lys	Met	Leu	Asn	His	Val	Leu	Asn	Ile	Cys	Glu	Lys	Asp	Gly	Thr
				95					100					105
Phe	Asp	Asn	Ile	Tyr	Leu	His	Val	Gln	Ile	Ser	Asn	Glu	Ser	Ala
				110					115					120
Ile	Asp	Phe	Tyr	Arg	Lys	Phe	Gly	Phe	Glu	Ile	Ile	Glu	Thr	Lys
				125					130					135
Lys	Asn	Tyr	Tyr	Lys	Arg	Ile	Glu	Pro	Ala	Asp	Ala	His	Val	Leu
				140					145					150
Gln	Lys	Asn	Leu	Lys	Val	Pro	Ser	Gly	Gln	Asn	Ala	Asp	Val	Gln
				155					160					165
Lys	Thr	Asp	Asn											

<210> 29  
 <211> 389  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2678733CD1

&lt;400&gt; 29.

```

Met Arg Val Leu Val Arg Arg Cys Trp Gly Pro Pro Leu Ala His
 1          5          10          15
Gly Ala Arg Arg Gly Arg Pro Ser Pro Gln Trp Arg Ala Leu Ala
 20          25          30
Arg Leu Gly Trp Glu Asp Cys Arg Asp Ser Arg Val Arg Glu Lys
 35          40          45
Pro Pro Trp Arg Val Leu Phe Phe Gly Thr Asp Gln Phe Ala Arg
 50          55          60
Glu Ala Leu Arg Ala Leu His Ala Ala Arg Glu Asn Lys Glu Glu
 65          70          75
Glu Leu Ile Asp Lys Leu Glu Val Val Thr Met Pro Ser Pro Ser
 80          85          90
Pro Lys Gly Leu Pro Val Lys Gln Tyr Ala Val Gln Ser Gln Leu
 95          100          105
Pro Val Tyr Glu Trp Pro Asp Val Gly Ser Gly Glu Tyr Asp Val
110          115          120
Gly Val Val Ala Ser Phe Gly Arg Leu Leu Asn Glu Ala Leu Ile
125          130          135
Leu Lys Phe Pro Tyr Gly Ile Leu Asn Val His Pro Ser Cys Leu
140          145          150
Pro Arg Trp Arg Gly Pro Ala Pro Val Ile His Thr Val Leu His
155          160          165
Gly Asp Thr Val Thr Gly Val Thr Ile Met Gln Ile Arg Pro Lys
170          175          180
Arg Phe Asp Val Gly Pro Ile Leu Lys Gln Glu Thr Val Pro Val
185          190          195
Pro Pro Lys Ser Thr Ala Lys Glu Leu Glu Ala Val Leu Ser Arg
200          205          210
Leu Gly Ala Asn Met Leu Ile Ser Val Leu Lys Asn Leu Pro Glu
215          220          225
Ser Leu Ser Asn Gly Arg Gln Gln Pro Met Glu Gly Ala Thr Tyr
230          235          240
Ala Pro Lys Ile Ser Ala Gly Thr Ser Cys Ile Lys Trp Glu Glu
245          250          255
Gln Thr Ser Glu Gln Ile Phe Arg Leu Tyr Arg Ala Ile Gly Asn
260          265          270
Ile Ile Pro Leu Gln Thr Leu Trp Met Ala Asn Thr Ile Lys Leu
275          280          285
Leu Asp Leu Val Glu Val Asn Ser Ser Val Leu Ala Asp Pro Lys
290          295          300
Leu Thr Gly Gln Ala Leu Ile Pro Gly Ser Val Ile Tyr His Lys
305          310          315
Gln Ser Gln Ile Leu Leu Val Tyr Cys Lys Asp Gly Trp Ile Gly
320          325          330
Val Arg Ser Val Met Leu Lys Lys Ser Leu Thr Ala Thr Asp Phe
335          340          345
Tyr Asn Gly Tyr Leu His Pro Trp Tyr Gln Lys Asn Ser Gln Ala
350          355          360
Gln Pro Ser Gln Cys Arg Phe Gln Thr Leu Arg Leu Pro Thr Lys
365          370          375
Lys Lys Gln Lys Lys Thr Val Ala Met Gln Gln Cys Ile Glu
380          385

```

&lt;210&gt; 30

&lt;211&gt; 600

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2768571CD1

&lt;400&gt; 30

```

Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val
 1          5          10          15
Gln Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala

```



	20		25		30									
Leu	Pro	Ser	Phe	Ile	Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His
	35								40					45
Gln	Arg	Thr	Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser	Leu	Ala
	50								55					60
Lys	Pro	Lys	Ser	Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr	Ile
	65								70					75
Tyr	Ala	Glu	Pro	Val	Pro	Glu	Asn	Asn	Ala	Leu	Asn	Thr	Gln	Thr
	80								85					90
Gln	Pro	Lys	Ala	His	Thr	Thr	Gly	Asp	Arg	Gly	Lys	Glu	Ala	Asn
	95								100					105
Gln	Ala	Pro	Pro	Glu	Glu	Gln	Asp	Lys	Val	Pro	His	Thr	Ala	Gln
	110								115					120
Arg	Ala	Ala	Trp	Lys	Ser	Pro	Glu	Lys	Glu	Lys	Thr	Met	Val	Asn
	125								130					135
Thr	Leu	Ser	Pro	Arg	Gly	Gln	Asp	Ala	Gly	Met	Ala	Ser	Gly	Arg
	140								145					150
Thr	Glu	Ala	Gln	Ser	Trp	Lys	Ser	Gln	Asp	Thr	Lys	Thr	Thr	Gln
	155								160					165
Gly	Asn	Gly	Gly	Gln	Thr	Arg	Lys	Leu	Thr	Ala	Ser	Arg	Thr	Val
	170								175					180
Ser	Glu	Lys	His	Gln	Gly	Lys	Ala	Ala	Thr	Thr	Ala	Lys	Thr	Leu
	185								190					195
Ile	Pro	Lys	Ser	Gln	His	Arg	Met	Leu	Ala	Pro	Thr	Gly	Ala	Val
	200								205					210
Ser	Thr	Arg	Thr	Arg	Gln	Lys	Gly	Val	Thr	Thr	Ala	Val	Ile	Pro
	215								220					225
Pro	Lys	Glu	Lys	Lys	Pro	Gln	Ala	Thr	Pro	Pro	Pro	Ala	Pro	Phe
	230								235					240
Gln	Ser	Pro	Thr	Thr	Gln	Arg	Asn	Gln	Arg	Leu	Lys	Ala	Ala	Asn
	245								250					255
Phe	Lys	Ser	Glu	Pro	Arg	Trp	Asp	Phe	Glu	Glu	Lys	Tyr	Ser	Phe
	260								265					270
Glu	Ile	Gly	Gly	Leu	Gln	Thr	Thr	Cys	Pro	Asp	Ser	Val	Lys	Ile
	275								280					285
Lys	Ala	Ser	Lys	Ser	Leu	Trp	Leu	Gln	Lys	Leu	Phe	Leu	Pro	Asn
	290								295					300
Leu	Thr	Leu	Phe	Leu	Asp	Ser	Arg	His	Phe	Asn	Gln	Ser	Glu	Trp
	305								310					315
Asp	Arg	Leu	Glu	His	Phe	Ala	Pro	Pro	Phe	Gly	Phe	Met	Glu	Leu
	320								325					330
Asn	Tyr	Ser	Leu	Val	Gln	Lys	Val	Val	Thr	Arg	Phe	Pro	Pro	Val
	335								340					345
Pro	Gln	Gln	Gln	Leu	Leu	Leu	Ala	Ser	Leu	Pro	Ala	Gly	Ser	Leu
	350								355					360
Arg	Cys	Ile	Thr	Cys	Ala	Val	Val	Gly	Asn	Gly	Gly	Ile	Leu	Asn
	365								370					375
Asn	Ser	His	Met	Gly	Gln	Glu	Ile	Asp	Ser	His	Asp	Tyr	Val	Phe
	380								385					390
Arg	Leu	Ser	Gly	Ala	Leu	Ile	Lys	Gly	Tyr	Glu	Gln	Asp	Val	Gly
	395								400					405
Thr	Arg	Thr	Ser	Phe	Tyr	Gly	Phe	Thr	Ala	Phe	Ser	Leu	Thr	Gln
	410								415					420
Ser	Leu	Leu	Ile	Leu	Gly	Asn	Arg	Gly	Phe	Lys	Asn	Val	Pro	Leu
	425								430					435
Gly	Lys	Asp	Val	Arg	Tyr	Leu	His	Phe	Leu	Glu	Gly	Thr	Arg	Asp
	440								445					450
Tyr	Glu	Trp	Leu	Glu	Ala	Leu	Leu	Met	Asn	Gln	Thr	Val	Met	Ser
	455								460					465
Lys	Asn	Leu	Phe	Trp	Phe	Arg	His	Arg	Pro	Gln	Glu	Ala	Phe	Arg
	470								475					480
Glu	Ala	Leu	His	Met	Asp	Arg	Tyr	Leu	Leu	Leu	His	Pro	Asp	Phe
	485								490					495
Leu	Arg	Tyr	Met	Lys	Asn	Arg	Phe	Leu	Arg	Ser	Lys	Thr	Leu	Asp
	500								505					510
Gly	Ala	His	Trp	Arg	Ile	Tyr	Arg	Pro	Thr	Thr	Gly	Ala	Leu	Leu
	515								520					525

Leu	Leu	Thr	Ala	Leu	Gln	Leu	Cys	Asp	Gln	Val	Ser	Ala	Tyr	Gly
				530					535					540
Phe	Ile	Thr	Glu	Gly	His	Glu	Arg	Phe	Ser	Asp	His	Tyr	Tyr	Asp
				545					550					555
Thr	Ser	Trp	Lys	Arg	Leu	Ile	Phe	Tyr	Ile	Asn	His	Asp	Phe	Lys
				560					565					570
Leu	Glu	Arg	Glu	Val	Trp	Lys	Arg	Leu	His	Asp	Glu	Gly	Ile	Ile
				575					580					585
Arg	Leu	Tyr	Gln	Arg	Pro	Gly	Pro	Gly	Thr	Ala	Lys	Ala	Lys	Asn
				590					595					600

<210> 31  
 <211> 448  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3189062CD1

<400> 31

Met	Arg	Glu	Asn	Val	Val	Val	Ser	Asn	Met	Glu	Arg	Glu	Ser	Gly
1				5					10					15
Lys	Pro	Val	Ala	Val	Val	Ala	Val	Val	Thr	Glu	Pro	Trp	Phe	Thr
				20					25					30
Gln	Arg	Tyr	Arg	Glu	Tyr	Leu	Gln	Arg	Gln	Lys	Leu	Phe	Asp	Thr
				35					40					45
Gln	His	Arg	Val	Glu	Lys	Met	Pro	Asp	Gly	Ser	Val	Ala	Leu	Pro
				50					55					60
Val	Leu	Gly	Glu	Thr	Leu	Pro	Glu	Gln	His	Leu	Gln	Glu	Leu	Arg
				65					70					75
Asn	Arg	Val	Ala	Pro	Gly	Ser	Pro	Cys	Met	Leu	Thr	Gln	Leu	Pro
				80					85					90
Asp	Pro	Val	Pro	Ser	Lys	Arg	Ala	Gln	Gly	Cys	Ser	Pro	Ala	Gln
				95					100					105
Lys	Leu	Cys	Leu	Glu	Val	Ser	Arg	Trp	Val	Val	Gly	Arg	Gly	Val
				110					115					120
Lys	Trp	Ser	Ala	Glu	Leu	Glu	Ala	Asp	Leu	Pro	Arg	Ser	Trp	Gln
				125					130					135
Arg	His	Gly	Asn	Leu	Leu	Leu	Leu	Ser	Glu	Asp	Cys	Phe	Gln	Ala
				140					145					150
Lys	Gln	Trp	Lys	Asn	Leu	Gly	Pro	Glu	Leu	Trp	Glu	Thr	Val	Ala
				155					160					165
Leu	Ala	Leu	Gly	Val	Gln	Arg	Leu	Ala	Lys	Arg	Gly	Arg	Val	Ser
				170					175					180
Pro	Asp	Gly	Thr	Arg	Thr	Pro	Ala	Val	Thr	Leu	Leu	Leu	Gly	Asp
				185					190					195
His	Gly	Trp	Val	Glu	His	Val	Asp	Asn	Gly	Ile	Arg	Tyr	Lys	Phe
				200					205					210
Asp	Val	Thr	Gln	Cys	Met	Phe	Ser	Phe	Gly	Asn	Ile	Thr	Glu	Lys
				215					220					225
Leu	Arg	Val	Ala	Ser	Leu	Ser	Cys	Ala	Gly	Glu	Val	Leu	Val	Asp
				230					235					240
Leu	Tyr	Ala	Gly	Ile	Gly	Tyr	Phe	Thr	Leu	Pro	Phe	Leu	Val	His
				245					250					255
Ala	Gly	Ala	Ala	Phe	Val	His	Ala	Cys	Glu	Trp	Asn	Pro	His	Ala
				260					265					270
Val	Val	Ala	Leu	Arg	Asn	Asn	Leu	Glu	Ile	Asn	Gly	Val	Ala	Asp
				275					280					285
Arg	Cys	Gln	Ile	His	Phe	Gly	Asp	Asn	Arg	Lys	Leu	Lys	Leu	Ser
				290					295					300
Asn	Ile	Ala	Asp	Arg	Val	Ile	Leu	Gly	Leu	Ile	Pro	Ser	Ser	Glu
				305					310					315
Glu	Gly	Trp	Pro	Ile	Ala	Cys	Gln	Val	Leu	Arg	Gln	Asp	Ala	Gly
				320					325					330
Gly	Ile	Leu	His	Ile	His	Gln	Asn	Val	Glu	Ser	Phe	Pro	Gly	Lys

Asn	Leu	Gln	Ala	Leu	Gly	Val	Ser	Lys	Val	Glu	Lys	Glu	His	Trp
				350					355					360
Leu	Tyr	Pro	Gln	Gln	Ile	Thr	Thr	Asn	Gln	Trp	Lys	Asn	Gly	Ala
				365					370					375
Thr	Arg	Asp	Ser	Arg	Gly	Lys	Met	Leu	Ser	Pro	Ala	Thr	Lys	Pro
				380					385					390
Glu	Trp	Gln	Arg	Trp	Ala	Glu	Ser	Ala	Glu	Thr	Arg	Ile	Ala	Thr
				395					400					405
Leu	Leu	Gln	Gln	Val	His	Gly	Lys	Pro	Trp	Lys	Thr	Gln	Ile	Leu
				410					415					420
His	Ile	Gln	Pro	Val	Lys	Ser	Tyr	Ala	Pro	His	Val	Asp	His	Ile
				425					430					435
Val	Leu	Asp	Leu	Glu	Cys	Cys	Pro	Cys	Pro	Ser	Val	Gly		
				440					445					

&lt;210&gt; 32

&lt;211&gt; 346

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3243884CD1

&lt;400&gt; 32

Met	Ala	Ala	Ser	Gly	Lys	Leu	Ser	Thr	Cys	Arg	Leu	Pro	Pro	Leu
1				5					10					15
Pro	Thr	Ile	Arg	Glu	Ile	Ile	Lys	Leu	Leu	Arg	Leu	Gln	Ala	Ala
				20					25					30
Lys	Gln	Leu	Ser	Gln	Asn	Phe	Leu	Leu	Asp	Leu	Arg	Leu	Thr	Asp
				35					40					45
Lys	Ile	Val	Arg	Lys	Ala	Gly	Asn	Leu	Thr	Asn	Ala	Tyr	Val	Tyr
				50					55					60
Glu	Val	Gly	Pro	Gly	Pro	Gly	Gly	Ile	Thr	Arg	Ser	Ile	Leu	Asn
				65					70					75
Ala	Asp	Val	Ala	Glu	Leu	Leu	Val	Val	Glu	Lys	Asp	Thr	Arg	Phe
				80					85					90
Ile	Pro	Gly	Leu	Gln	Met	Leu	Ser	Asp	Ala	Ala	Pro	Gly	Lys	Leu
				95					100					105
Arg	Ile	Val	His	Gly	Asp	Val	Leu	Thr	Phe	Lys	Val	Glu	Lys	Ala
				110					115					120
Phe	Ser	Glu	Ser	Leu	Lys	Arg	Pro	Trp	Glu	Asp	Asp	Pro	Pro	Asn
				125					130					135
Val	His	Ile	Ile	Gly	Asn	Leu	Pro	Phe	Ser	Val	Ser	Thr	Pro	Leu
				140					145					150
Ile	Ile	Lys	Trp	Leu	Glu	Asn	Ile	Ser	Cys	Arg	Asp	Gly	Pro	Phe
				155					160					165
Val	Tyr	Gly	Arg	Thr	Gln	Met	Thr	Leu	Thr	Phe	Gln	Lys	Glu	Val
				170					175					180
Ala	Glu	Arg	Leu	Ala	Ala	Asn	Thr	Gly	Ser	Lys	Gln	Arg	Ser	Arg
				185					190					195
Leu	Ser	Val	Met	Ala	Gln	Tyr	Leu	Cys	Asn	Val	Arg	His	Ile	Phe
				200					205					210
Thr	Ile	Pro	Gly	Gln	Ala	Phe	Val	Pro	Lys	Pro	Glu	Val	Asp	Val
				215					220					225
Gly	Val	Val	His	Phe	Thr	Pro	Leu	Ile	Gln	Pro	Lys	Ile	Glu	Gln
				230					235					240
Pro	Phe	Lys	Leu	Val	Glu	Lys	Val	Val	Gln	Asn	Val	Phe	Gln	Phe
				245					250					255
Arg	Arg	Lys	Tyr	Cys	His	Arg	Gly	Leu	Arg	Met	Leu	Phe	Pro	Glu
				260					265					270
Ala	Gln	Arg	Leu	Glu	Ser	Thr	Gly	Arg	Leu	Leu	Glu	Leu	Ala	Asp
				275					280					285
Ile	Asp	Pro	Thr	Leu	Arg	Pro	Arg	Gln	Leu	Ser	Ile	Ser	His	Phe
				290					295					300
Lys	Ser	Leu	Cys	Asp	Val	Tyr	Arg	Lys	Met	Cys	Asp	Glu	Asp	Pro

	305		310		315
Gln Leu Phe Ala Tyr Asn Phe Arg Glu Glu Leu Lys Arg Arg Lys					
	320		325		330
Ser Lys Asn Glu Glu Lys Glu Glu Asp Asp Ala Glu Asn Tyr Arg					
	335		340		345

Leu

<210> 33  
 <211> 173  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3400578CD1

<400> 33

Met Ala Ser Ile Leu Arg Thr Pro Gln Ala Leu Gln Leu Thr Leu		
1 5 10 15		
Ala Leu Ile Lys Pro Asp Ala Val Ala His Pro Leu Ile Leu Glu		
20 25 30		
Ala Val His Gln Gln Ile Leu Ser Asn Lys Phe Leu Ile Val Arg		
35 40 45		
Met Arg Glu Leu Leu Trp Arg Lys Glu Asp Cys Gln Arg Phe Tyr		
50 55 60		
Arg Glu His Glu Ala Gly Pro Ile Arg Ala Tyr Ile Leu Ala His		
65 70 75		
Lys Asp Ala Ile Gln Leu Trp Arg Thr Leu Met Gly Pro Thr Arg		
80 85 90		
Val Phe Arg Ala Arg His Val Ala Pro Asp Ser Ile Arg Gly Ser		
95 100 105		
Phe Gly Leu Thr Asp Thr Arg Asn Thr Thr His Gly Ser Asp Ser		
110 115 120		
Val Val Ser Ala Ser Arg Glu Ile Ala Ala Phe Phe Pro Asp Phe		
125 130 135		
Ser Glu Gln Arg Trp Tyr Glu Glu Glu Pro Gln Leu Arg Cys		
140 145 150		
Gly Pro Val Cys Tyr Ser Pro Glu Gly Gly Val His Tyr Val Ala		
155 160 165		
Gly Thr Gly Gly Leu Gly Pro Ala		
170		

<210> 34  
 <211> 445  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3422577CD1

<400> 34

Met Thr Glu Leu Arg Gln Arg Val Ala His Glu Pro Val Ala Pro		
1 5 10 15		
Pro Glu Asp Lys Glu Ser Glu Ser Glu Ala Lys Val Asp Gly Glu		
20 25 30		
Thr Ala Ser Asp Ser Glu Ser Arg Ala Glu Ser Ala Pro Leu Pro		
35 40 45		
Val Ser Ala Asp Asp Thr Pro Glu Val Leu Asn Arg Ala Leu Ser		
50 55 60		
Asn Leu Ser Ser Arg Trp Lys Asn Trp Trp Val Arg Gly Ile Leu		
65 70 75		
Thr Leu Ala Met Ile Ala Phe Phe Phe Ile Ile Ile Tyr Leu Gly		
80 85 90		
Pro Met Val Leu Met Ile Ile Val Met Cys Val Gln Ile Lys Cys		
95 100 105		

Phe	His	Glu	Ile	Ile	Thr	Ile	Gly	Tyr	Asn	Val	Tyr	His	Ser	Tyr
				110					115					120
Asp	Leu	Pro	Trp	Phe	Arg	Thr	Leu	Ser	Trp	Tyr	Phe	Leu	Leu	Cys
				125					130					135
Val	Asn	Tyr	Phe	Phe	Tyr	Gly	Glu	Thr	Val	Thr	Asp	Tyr	Phe	Phe
				140					145					150
Thr	Leu	Val	Gln	Arg	Glu	Glu	Pro	Leu	Arg	Ile	Leu	Ser	Lys	Tyr
				155					160					165
His	Arg	Phe	Ile	Ser	Phe	Thr	Leu	Tyr	Leu	Ile	Gly	Phe	Cys	Met
				170					175					180
Phe	Val	Leu	Ser	Leu	Val	Lys	Lys	His	Tyr	Arg	Leu	Gln	Phe	Tyr
				185					190					195
Met	Phe	Gly	Trp	Thr	His	Val	Thr	Leu	Leu	Ile	Val	Val	Thr	Gln
				200					205					210
Ser	His	Leu	Val	Ile	His	Asn	Leu	Phe	Glu	Gly	Met	Ile	Trp	Phe
				215					220					225
Ile	Val	Pro	Ile	Ser	Cys	Val	Ile	Cys	Asn	Asp	Ile	Met	Ala	Tyr
				230					235					240
Met	Phe	Gly	Phe	Phe	Phe	Gly	Arg	Thr	Pro	Leu	Ile	Lys	Leu	Ser
				245					250					255
Pro	Lys	Lys	Thr	Trp	Glu	Gly	Phe	Ile	Gly	Gly	Phe	Phe	Ala	Thr
				260					265					270
Val	Val	Phe	Gly	Leu	Leu	Leu	Ser	Tyr	Val	Met	Ser	Gly	Tyr	Arg
				275					280					285
Cys	Phe	Val	Cys	Pro	Val	Glu	Tyr	Asn	Asn	Asp	Thr	Asn	Ser	Phe
				290					295					300
Thr	Val	Asp	Cys	Glu	Pro	Ser	Asp	Leu	Phe	Arg	Leu	Gln	Glu	Tyr
				305					310					315
Asn	Ile	Pro	Gly	Val	Ile	Gln	Ser	Val	Ile	Gly	Trp	Lys	Thr	Val
				320					325					330
Arg	Met	Tyr	Pro	Phe	Gln	Ile	His	Ser	Ile	Ala	Leu	Ser	Thr	Phe
				335					340					345
Ala	Ser	Leu	Ile	Gly	Pro	Phe	Gly	Gly	Phe	Phe	Ala	Ser	Gly	Phe
				350					355					360
Lys	Arg	Ala	Phe	Lys	Ile	Lys	Asp	Phe	Ala	Asn	Thr	Ile	Pro	Gly
				365					370					375
His	Gly	Gly	Ile	Met	Asp	Arg	Phe	Asp	Cys	Gln	Tyr	Leu	Met	Ala
				380					385					390
Thr	Phe	Val	Asn	Val	Tyr	Ile	Ala	Ser	Phe	Ile	Arg	Gly	Pro	Asn
				395					400					405
Pro	Ser	Lys	Leu	Ile	Gln	Gln	Phe	Leu	Thr	Leu	Arg	Pro	Asp	Gln
				410					415					420
Gln	Leu	His	Ile	Phe	Asn	Thr	Leu	Arg	Ser	His	Leu	Ile	Asp	Lys
				425					430					435
Gly	Met	Leu	Thr	Ser	Thr	Thr	Glu	Asp	Glu					
				440					445					

&lt;210&gt; 35

&lt;211&gt; 420

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3706809CD1

&lt;400&gt; 35

Met	Ala	Ala	Leu	Val	Arg	Pro	Ala	Arg	Phe	Val	Val	Arg	Pro	Leu
1				5					10					15
Leu	Gln	Val	Val	Gln	Ala	Trp	Asp	Leu	Asp	Ala	Arg	Arg	Trp	Val
				20					25					30
Arg	Ala	Leu	Arg	Arg	Ser	Pro	Val	Lys	Val	Val	Phe	Pro	Ser	Gly
				35					40					45
Glu	Val	Val	Glu	Gln	Lys	Arg	Ala	Pro	Gly	Lys	Gln	Pro	Arg	Lys
				50					55					60
Ala	Pro	Ser	Glu	Ala	Ser	Ala	Gln	Glu	Gln	Arg	Glu	Lys	Gln	Pro
				65					70					75

Leu	Glu	Glu	Ser	Ala	Ser	Arg	Ala	Pro	Ser	Thr	Trp	Glu	Glu	Ser	
				80					85					90	
Gly	Leu	Arg	Tyr	Asp	Lys	Ala	Tyr	Pro	Gly	Asp	Arg	Arg	Leu	Ser	
				95					100					105	
Ser	Val	Met	Thr	Ile	Val	Lys	Ser	Arg	Pro	Phe	Arg	Glu	Lys	Gln	
				110					115					120	
Gly	Lys	Ile	Leu	Leu	Glu	Gly	Arg	Arg	Leu	Ile	Ser	Asp	Ala	Leu	
				125					130					135	
Lys	Ala	Gly	Ala	Val	Pro	Lys	Met	Phe	Phe	Phe	Ser	Arg	Leu	Glu	
				140					145					150	
Tyr	Leu	Lys	Glu	Leu	Pro	Val	Asp	Lys	Leu	Lys	Gly	Val	Ser	Leu	
				155					160					165	
Ile	Lys	Val	Lys	Phe	Glu	Asp	Ile	Lys	Asp	Trp	Ser	Asp	Leu	Val	
				170					175					180	
Thr	Pro	Gln	Gly	Ile	Met	Gly	Ile	Phe	Ala	Lys	Pro	Asp	His	Val	
				185					190					195	
Lys	Met	Thr	Tyr	Pro	Lys	Thr	Gln	Leu	Gln	His	Ser	Leu	Pro	Leu	
				200					205					210	
Leu	Leu	Ile	Cys	Asp	Asn	Leu	Arg	Asp	Pro	Gly	Asn	Leu	Gly	Thr	
				215					220					225	
Ile	Leu	Arg	Ser	Ala	Ala	Gly	Ala	Gly	Cys	Ser	Lys	Val	Leu	Leu	
				230					235					240	
Thr	Lys	Gly	Cys	Val	Asp	Ala	Trp	Glu	Pro	Lys	Val	Leu	Arg	Ala	
				245					250					255	
Gly	Met	Gly	Ala	His	Phe	Arg	Met	Pro	Ile	Ile	Asn	Asn	Leu	Glu	
				260					265					270	
Trp	Glu	Thr	Val	Pro	Asn	Tyr	Leu	Pro	Pro	Asp	Thr	Arg	Val	Tyr	
				275					280					285	
Val	Ala	Asp	Asn	Cys	Gly	Leu	Tyr	Ala	Gln	Ala	Glu	Met	Ser	Asn	
				290					295					300	
Lys	Ala	Ser	Asp	His	Gly	Trp	Val	Cys	Asp	Gln	Arg	Val	Met	Lys	
				305					310					315	
Phe	His	Lys	Tyr	Glu	Glu	Glu	Glu	Asp	Val	Glu	Thr	Gly	Ala	Ser	
				320					325					330	
Gln	Asp	Trp	Leu	Pro	His	Val	Glu	Val	Gln	Ser	Tyr	Asp	Ser	Asp	
				335					340					345	
Trp	Thr	Glu	Ala	Pro	Ala	Ala	Val	Val	Ile	Gly	Gly	Glu	Thr	Tyr	
				350					355					360	
Gly	Val	Ser	Leu	Glu	Ser	Leu	Gln	Leu	Ala	Glu	Ser	Thr	Gly	Gly	
				365					370					375	
Lys	Arg	Leu	Leu	Ile	Pro	Val	Val	Pro	Gly	Val	Asp	Ser	Leu	Asn	
				380					385					390	
Ser	Ala	Met	Ala	Ala	Ser	Ile	Leu	Leu	Phe	Glu	Gly	Lys	Arg	Gln	
				395					400					405	
Leu	Arg	Gly	Arg	Ala	Glu	Asp	Leu	Ser	Arg	Asp	Arg	Ser	Tyr	His	
				410					415					420	

&lt;210&gt; 36

&lt;211&gt; 354

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3745914CD1

&lt;400&gt; 36

Met	Ala	Pro	Ala	Lys	Ala	Thr	Asn	Val	Val	Arg	Leu	Leu	Leu	Gly	
				5					10					15	
Ser	Thr	Ala	Leu	Trp	Leu	Ser	Gln	Leu	Gly	Ser	Gly	Thr	Val	Ala	
				20					25					30	
Ala	Ser	Lys	Ser	Val	Thr	Ala	His	Leu	Ala	Ala	Lys	Trp	Pro	Glu	
				35					40					45	
Thr	Pro	Leu	Leu	Leu	Glu	Ala	Ser	Glu	Phe	Met	Ala	Glu	Glu	Ser	
				50					55					60	
Asn	Glu	Lys	Phe	Trp	Gln	Phe	Leu	Glu	Thr	Val	Gln	Glu	Leu	Ala	

Ile Tyr Lys Gln Thr	65	Glu Ser Asp Tyr	70	Ser Tyr Tyr Asn Leu	75
Leu Lys Lys Ala Gly	80	Gln Phe Leu Asp	85	Asn Leu His Ile Asn	90
Leu Lys Phe Ala Phe	95	Ser Ile Arg Ala	100	Tyr Ser Pro Ala Ile	105
Met Phe Gln Gln Ile	110	Ala Ala Asp Glu	115	Pro Pro Pro Asp Gly	120
Asn Ala Phe Val Val	125	Ile His Lys Lys	130	His Thr Cys Lys Ile	135
Glu Ile Lys Lys Leu	140	Leu Lys Lys Ala	145	Ala Ser Arg Thr Arg	150
Tyr Leu Phe Lys Gly	155	Asp His Lys Phe	160	Pro Thr Asn Lys Glu	165
Leu Pro Val Val Ile	170	Leu Tyr Ala Glu	175	Met Gly Thr Arg Thr	180
Ser Ala Phe His Lys	185	Val Leu Ser Glu	190	Lys Ala Gln Asn Glu	195
Ile Leu Tyr Val Leu	200	Arg His Tyr Ile	205	Gln Lys Pro Ser Ser	210
Lys Met Tyr Leu Ser	215	Gly Tyr Gly Val	220	Glu Leu Ala Ile Lys	225
Thr Glu Tyr Lys Ala	230	Leu Asp Asp Thr	235	Gln Val Lys Thr Val	240
Asn Thr Thr Val Glu	245	Asp Glu Thr Glu	250	Thr Asn Glu Val Gln	255
Phe Leu Phe Gly Lys	260	Leu Lys Glu Ile	265	Tyr Ser Asp Leu Arg	270
Asn Leu Thr Ala Phe	275	Gln Lys Tyr Leu	280	Ile Glu Ser Asn Lys	285
Met Met Pro Leu Lys	290	Val Trp Glu Leu	295	Gln Asp Leu Ser Phe	300
Ala Ala Ser Gln Ile	305	Met Ser Ala Pro	310	Val Tyr Asp Ala Ile	315
Leu Met Lys Asp Ile	320	Ser Gln Asn Phe	325	Pro Ile Lys Ala Arg	330
Gln Met Ile Gly Asn	335	Val Leu Ile Gly	340		345
	350				

<210> 37  
 <211> 198  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4000776CD1

---

<400> 37

Met Ser Ser Lys Arg	1	Ser His Tyr Asp	10	Ser Ala Leu Lys Arg	15
Val Ile Val Tyr Ala	5	Glu Lys His Gly	20	Asn Arg Ala Ala Gly	30
Thr Phe Asp Ile Ser	25	Glu Ala Asn Ile	35	Arg Arg Trp Arg Asn	45
Arg Asn Ser Ile Phe	40	Ser Cys Lys Ala	50	Thr Thr Lys Cys Phe	60
Gly Pro Lys Lys Gly	55	Arg Tyr Pro Gln	65	Val Asp Glu Ala Val	75
Arg Phe Val Ser Glu	70	Thr Arg Ala Lys	80	Gly Leu Pro Ile Thr	90
Gln Ala Met Gln Leu	85	Lys Ala Gly Glu	95	Val Ala Lys Thr Leu	105
Ile Asp Glu Thr Lys	100	Phe Lys Ala Thr	110	Arg Gly Trp Cys Asp	120
Phe Met Arg Arg Ala	115	Gly Leu Ser Leu	125	Arg His Gln Thr Ser	135

Cys	Pro	Lys	Leu	125	Pro	Thr	Ala	Ile	Lys	130	Gln	Lys	Thr	Val	Leu	135
His	Ser	Phe	Lys	140	Lys	Cys	Cys	Ile	Thr	145	Ser	Thr	Leu	Asp	Asn	150
Gly	Arg	Asp	Val	155	Leu	Trp	Lys	Asn	Ala	160	Asp	Ile	Asn	Asp	Cys	165
Leu	Lys	Ser	Asp	170	Ser	Glu	Glu	Leu	Asp	175	Ser	Glu	Tyr	Glu	Val	180
Ile	Ile	Thr		185						190						195

<210> 38  
 <211> 296  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4071304CD1

<400> 38

Met	Met	Leu	Pro	Leu	5	Gln	Gly	Ala	Gln	10	Met	Leu	Gln	Met	Leu	15
Lys	Ser	Leu	Arg	Lys	20	Ser	Leu	Pro	Ala	25	Ser	Leu	Lys	Val	Tyr	30
Thr	Val	Phe	His	Ile	35	Asn	His	Gly	Asn	40	Pro	Phe	Asn	Leu	Lys	45
Val	Val	Asp	Lys	Trp	50	Pro	Asp	Phe	Asn	55	Thr	Val	Val	Val	Cys	60
Gln	Glu	Gln	Asp	Met	65	Thr	Asp	Asp	Leu	70	Asp	His	Tyr	Thr	Asn	75
Tyr	Gln	Ile	Tyr	Ser	80	Lys	Asp	Pro	Gln	85	Asn	Cys	Gln	Glu	Phe	90
Gly	Ser	Pro	Glu	Leu	95	Ile	Asn	Trp	Lys	100	Gln	His	Leu	Gln	Ile	105
Ser	Ser	Gln	Pro	Ser	110	Leu	Asn	Glu	Ala	115	Ile	Gln	Asn	Leu	Ala	120
Ile	Lys	Ser	Phe	Lys	125	Val	Lys	Gln	Thr	130	Gln	Arg	Ile	Leu	Tyr	135
Ala	Ala	Glu	Thr	Ala	140	Lys	Glu	Leu	Thr	145	Pro	Phe	Leu	Leu	Lys	150
Lys	Ile	Leu	Ser	Pro	155	Ser	Gly	Gly	Lys	160	Pro	Lys	Ala	Ile	Asn	165
Glu	Met	Phe	Lys	Leu	170	Ser	Ser	Met	Asp	175	Val	Thr	His	Ala	His	180
Val	Asn	Lys	Phe	Trp	185	His	Phe	Gly	Gly	190	Asn	Glu	Arg	Ser	Gln	195
Phe	Ile	Glu	Arg	Cys	200	Ile	Gln	Thr	Phe	205	Pro	Thr	Cys	Cys	Leu	210
Gly	Pro	Glu	Gly	Thr	215	Pro	Val	Cys	Trp	220	Asp	Leu	Met	Asp	Gln	225
Gly	Glu	Met	Arg	Met	230	Ala	Gly	Thr	Phe	235	Ala	Glu	Tyr	Arg	Leu	240
Gly	Leu	Val	Thr	Tyr	245	Val	Ile	Tyr	Ser	250	His	Ala	Gln	Lys	Leu	255
Lys	Leu	Gly	Phe	Pro	260	Val	Tyr	Ser	His	265	Val	Asp	Tyr	Ser	Asn	270
Ala	Met	Gln	Lys	Met	275	Ser	Tyr	Thr	Leu	280	Gln	His	Val	Pro	Ile	285
Arg	Ser	Trp	Asn	Gln	290	Trp	Asn	Cys	Val	295	Pro	Leu				

<210> 39  
 <211> 214  
 <212> PRT  
 <213> Homo sapiens



<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4344970CD1

<400> 39

Met	Ala	Gly	Glu	Asn	Phe	Ala	Thr	Pro	Phe	His	Gly	His	Val	Gly	
1				5					10					15	
Arg	Gly	Ala	Phe	Ser	Asp	Val	Tyr	Glu	Pro	Ala	Glu	Asp	Thr	Phe	
				20					25					30	
Leu	Leu	Leu	Asp	Ala	Leu	Glu	Ala	Ala	Ala	Ala	Glu	Leu	Ala	Gly	
				35					40					45	
Val	Glu	Ile	Cys	Leu	Glu	Val	Gly	Ser	Gly	Ser	Gly	Val	Val	Ser	
				50					55					60	
Ala	Phe	Leu	Ala	Ser	Met	Ile	Gly	Pro	Gln	Ala	Leu	Tyr	Met	Cys	
				65					70					75	
Thr	Asp	Ile	Asn	Pro	Glu	Ala	Ala	Ala	Cys	Thr	Leu	Glu	Thr	Ala	
				80					85					90	
Arg	Cys	Asn	Lys	Val	His	Ile	Gln	Pro	Val	Ile	Thr	Asp	Leu	Val	
				95					100					105	
Lys	Gly	Leu	Leu	Pro	Arg	Leu	Thr	Glu	Lys	Val	Asp	Leu	Leu	Val	
				110					115					120	
Phe	Asn	Pro	Pro	Tyr	Val	Val	Thr	Pro	Pro	Gln	Glu	Val	Gly	Ser	
				125					130					135	
His	Gly	Ile	Glu	Ala	Ala	Trp	Ala	Gly	Gly	Arg	Asn	Gly	Arg	Glu	
				140					145					150	
Val	Met	Asp	Arg	Phe	Phe	Pro	Leu	Val	Pro	Asp	Leu	Leu	Ser	Pro	
				155					160					165	
Arg	Gly	Leu	Phe	Tyr	Leu	Val	Thr	Ile	Lys	Glu	Asn	Asn	Pro	Glu	
				170					175					180	
Glu	Ile	Leu	Lys	Ile	Met	Lys	Thr	Lys	Gly	Leu	Gln	Gly	Thr	Thr	
				185					190					195	
Ala	Leu	Ser	Arg	Gln	Ala	Gly	Gln	Glu	Thr	Leu	Ser	Val	Leu	Lys	
				200					205					210	
Phe	Thr	Lys	Ser												

<210> 40  
 <211> 322  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5392302CD1

<400> 40

Met	Ala	Ala	Ser	Gly	Glu	Pro	Gln	Arg	Gln	Trp	Gln	Glu	Glu	Val	
1				5					10					15	
Ala	Ala	Val	Val	Val	Val	Gly	Ser	Cys	Met	Thr	Asp	Leu	Val	Ser	
				20					25					30	
Leu	Thr	Ser	Arg	Leu	Pro	Lys	Thr	Gly	Glu	Thr	Ile	His	Gly	His	
				35					40					45	
Lys	Phe	Phe	Ile	Gly	Phe	Gly	Gly	Lys	Gly	Ala	Asn	Gln	Cys	Val	
				50					55					60	
Gln	Ala	Ala	Arg	Leu	Gly	Ala	Met	Thr	Ser	Met	Val	Cys	Lys	Val	
				65					70					75	
Gly	Lys	Asp	Ser	Phe	Gly	Asn	Asp	Tyr	Ile	Glu	Asn	Leu	Lys	Gln	
				80					85					90	
Asn	Asp	Ile	Ser	Thr	Glu	Phe	Thr	Tyr	Gln	Thr	Lys	Asp	Ala	Ala	
				95					100					105	
Thr	Gly	Thr	Ala	Ser	Ile	Ile	Val	Asn	Asn	Glu	Gly	Gln	Asn	Ile	
				110					115					120	
Ile	Val	Ile	Val	Ala	Gly	Ala	Asn	Leu	Leu	Leu	Asn	Thr	Glu	Asp	
				125					130					135	
Leu	Arg	Ala	Ala	Ala	Asn	Val	Ile	Ser	Arg	Ala	Lys	Val	Met	Val	
				140					145					150	
Cys	Gln	Leu	Glu	Ile	Thr	Pro	Ala	Thr	Ser	Leu	Glu	Ala	Leu	Thr	

155	160	165
Met Ala Arg Arg Ser Gly Val Lys Thr	Leu Phe Asn Pro Ala Pro	
170	175	180
Ala Ile Ala Asp Leu Asp Pro Gln Phe Tyr Thr Leu Ser Asp Val		
185	190	195
Phe Cys Cys Asn Glu Ser Glu Ala Glu Ile Leu Thr Gly Leu Thr		
200	205	210
Val Gly Ser Ala Ala Asp Ala Gly Glu Ala Ala Leu Val Leu Leu		
215	220	225
Lys Arg Gly Cys Gln Val Val Ile Ile Thr Leu Gly Ala Glu Gly		
230	235	240
Cys Val Val Leu Ser Gln Thr Glu Pro Glu Pro Lys His Ile Pro		
245	250	255
Thr Glu Lys Val Lys Ala Val Asp Thr Thr Gly Ala Gly Asp Ser		
260	265	270
Phe Val Gly Ala Leu Ala Phe Tyr Leu Ala Tyr Tyr Pro Asn Leu		
275	280	285
Ser Leu Glu Asp Met Leu Asn Arg Ser Asn Phe Ile Ala Ala Val		
290	295	300
Ser Val Gln Ala Ala Gly Thr Gln Ser Ser Tyr Pro Tyr Lys Lys		
305	310	315
Asp Leu Pro Leu Thr Leu Phe		
320		

<210> 41  
 <211> 87  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5555235CD1

<400> 41	
Met Ser Thr Ser Val Pro Gln Gly His Thr Trp Thr Gln Arg Val	
1 5	10 15
Lys Lys Asp Asp Glu Glu Asp Pro Leu Asp Gln Leu Ile Ser	
20 25	30
Arg Ser Gly Cys Ala Ala Ser His Phe Ala Val Gln Glu Cys Met	
35 40	45
Ala Gln His Gln Asp Trp Arg Gln Cys Gln Pro Gln Val Gln Ala	
50 55	60
Phe Lys Asp Cys Met Ser Glu Gln Gln Ala Arg Arg Gln Glu Glu	
65 70	75
Leu Gln Arg Arg Gln Glu Gln Ala Gly Ala His His	
80 85	

<210> 42  
 <211> 378  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5573296CD1

<400> 42	
Met Asp Leu Ala Gly Leu Leu Lys Ser Gln Phe Leu Cys His Leu	
1 5	10 15
Val Phe Cys Tyr Val Phe Ile Ala Ser Gly Leu Ile Ile Asn Thr	
20 25	30
Ile Gln Leu Phe Thr Leu Leu Leu Trp Pro Ile Asn Lys Gln Leu	
35 40	45
Phe Arg Lys Ile Asn Cys Arg Leu Ser Tyr Cys Ile Ser Ser Gln	
50 55	60
Leu Val Met Leu Leu Glu Trp Trp Ser Gly Thr Glu Cys Thr Ile	
65 70	75

Phe Thr Asp Pro Arg Ala Tyr Leu Lys Tyr Gly Lys Glu Asn Ala  
 80 85 90  
 Ile Val Val Leu Asn His Lys Phe Glu Ile Asp Phe Leu Cys Gly  
 95 100 105  
 Trp Ser Leu Ser Glu Arg Phe Gly Leu Leu Gly Gly Ser Lys Val  
 110 115 120  
 Leu Ala Lys Lys Glu Leu Ala Tyr Val Pro Ile Ile Gly Trp Met  
 125 130 135  
 Trp Tyr Phe Thr Glu Met Val Phe Cys Ser Arg Lys Trp Glu Gln  
 140 145 150  
 Asp Arg Lys Thr Val Ala Thr Ser Leu Gln His Leu Arg Asp Tyr  
 155 160 165  
 Pro Glu Lys Tyr Phe Phe Leu Ile His Cys Glu Gly Thr Arg Phe  
 170 175 180  
 Thr Glu Lys Lys His Glu Ile Ser Met Gln Val Ala Arg Ala Lys  
 185 190 195  
 Gly Leu Pro Arg Leu Lys His His Leu Leu Pro Arg Thr Lys Gly  
 200 205 210  
 Phe Ala Ile Thr Val Arg Ser Leu Arg Asn Val Val Ser Ala Val  
 215 220 225  
 Tyr Asp Cys Thr Leu Asn Phe Arg Asn Asn Glu Asn Pro Thr Leu  
 230 235 240  
 Leu Gly Val Leu Asn Gly Lys Lys Tyr His Ala Asp Leu Tyr Val  
 245 250 255  
 Arg Arg Ile Pro Leu Glu Asp Ile Pro Glu Asp Asp Asp Glu Cys  
 260 265 270  
 Ser Ala Trp Leu His Lys Leu Tyr Gln Glu Lys Asp Ala Phe Gln  
 275 280 285  
 Glu Glu Tyr Tyr Arg Thr Gly Thr Phe Pro Glu Thr Pro Met Val  
 290 295 300  
 Pro Pro Arg Arg Pro Trp Thr Leu Val Asn Trp Leu Phe Trp Ala  
 305 310 315  
 Ser Leu Val Leu Tyr Pro Phe Phe Gln Phe Leu Val Ser Met Ile  
 320 325 330  
 Arg Ser Gly Ser Ser Leu Thr Leu Ala Ser Phe Ile Leu Val Phe  
 335 340 345  
 Phe Val Ala Ser Val Gly Val Arg Trp Met Ile Gly Val Thr Glu  
 350 355 360  
 Ile Asp Lys Gly Ser Ala Tyr Gly Asn Ser Asp Ser Lys Gln Lys  
 365 370 375  
 Leu Asn Asp

&lt;210&gt; 43

&lt;211&gt; 1322

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 016233CB1

&lt;400&gt; 43

gcagcatcac gctgacctct gcctgggatg taaaccggac cagccgctgc gggcaaagga 60  
 aggtctttgg ctcttcggg aaaccagcc ccgtaccgg gctccgagcg gctcgaggc 120  
 gagcgacagc gacctagcc ccggcagcgc ccagcggcgg ctgcggaaag cggagggagt 180  
 ccgacgcggg cgcgggcggg gagcgtgcgt ccgttcgcac aggcagcggg aggagggcg 240  
 gcgcgaacca tggccgggga cagcgagcag accctgcaga accaccagca gcccaacggc 300  
 ggcgagccct tccttatagg cgtcagcggg ggaacagcta gcggcaagtc ttccgtgtgt 360  
 gctaagatcg tgcagtcct ggggcagaat gaggtggact atcgccagaa gcaggtgggtc 420  
 atctgagcc aggatagctt ctaccgtgtc cttacctcg agcagaaggc caaagccctg 480  
 aagggccagt tcaactttga ccaccggat gcctttgaca atgaactcat tctcaaaaca 540  
 ctcaaagaaa tcactgaagg gaaaacagtc cagatccccg tgtatgactt tgtctcccat 600  
 tcccgaagg aggagacagt tactgtctat cccgcagacg tgggtgctct tgaagggatc 660  
 ctggccttct actcccagga ggtacgagac ctgttccaga tgaagctttt tgtggatata 720  
 gatgcggaca cccggctctc acgcagagta ttaagggaca tcagcgagag aggcagggat 780  
 cttgagcaga ttttatctca gtacattacg ttcgtcaagc ctgcctttga ggaattctgc 840

```

ttgccaacaa agaagtatgc tgatgtgatc atccctagag gtgcagataa tctgggtggcc 900
atcaacctca tctgtgcagca catccaggac atcctgaatg gagggccctc caaacggcag 960
accaatggct gtctcaacgg ctacacccct tcacgcaaga ggcaggcatc ggagtccagc 1020
agcaggccgc attgacccgt ctccatcgga ccccgagccc tatctccaag agacagagga 1080
ggggtcagga ggcactgtct atctgtacat actgtttcct atgacattac tgtattttaag 1140
aaaacacccat ggagatgaaa tgcctttgat ttttttttct tttttgtact ttggaacgac 1200
aaaatgaaac agaacttgac cctgagctta aataacaaaa ctgtgccaac tactactggc 1260
gatgcctaata tatgaatcca acgtgtaacc agttataaat acatatatat ataaaaaaaa 1320
aa

```

&lt;210&gt; 44

&lt;211&gt; 1302

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 078336CB1

&lt;400&gt; 44

```

aagttttgggt gtttaaatgt tgtgttctat ctcttcact acaaacagga tacgcattgg 60
tgtgggtatct acttagcaac agaaataaag aaattagatc agtgccagtt tgaactcatc 120
atccagtggc tttctaaatc tgtaggcaag actcttaata aataaatagt caagccactt 180
gaaatcaggg atccaatact cttctcccat agcaaaatca agagctgttt aagctctaag 240
gctctatcta gcccaaaaac aaatgctata atgttttact tgggtgggtgt tctaaattca 300
gggtgatatcc aagaactgta tgacaccacc ttggccctgg gccacgaggc ggcttttctca 360
gatgactgcg atttgcctc tgctcaggac ataaacagac tctgtgggact tcagaacaca 420
tatatgggct atctggacta ccggaagaag gccatcaagg accttggcat cagccccagc 480
acctgctctt tcaatcctgg tgtgattgtt gccaacatga cagaatggaa gcaccagcgc 540
atcaccaagc aattggagaa atggatgcaa aagaatgtgg aggaaaacct ctatagcagc 600
tccctgggag gaggggtggc cacctcccca atgctgattg tgtttcatgg gaaatattcc 660
acaattaacc cctgtggca cataaggcac ctgggctgga atccagatgc cagatattcg 720
gagcattttc tgcaggaagc taaattact cactggaatg gaagacataa accttgggac 780
ttcctagtg ttcacaacga cttatgggaa agctgggttg ttctgacctc tgcagggata 840
tttaaactca atcaccatag ctgatataac tctaccctta aaatattccc tgtatagaaa 900
tgttgaattg tccctttgta gcccaactata acattgttct ttatgaatat tacctttgat 960
acatatgac cacaatataa aaacaaaaaa ctactgtgtg caaattatac cttggaccat 1020
ataggcattg ataaacttct ttaagtacat gtgataacta tggaaatcaa gattatgtga 1080
ctgaaaaaca taaaggaaga gacccatcta gataacagca atcaacctgc ttaattctga 1140
atgacaatta tatccacaaa tttttaaaac ttctacatgt atttttcaca tgaagatctc 1200
cttaacaggt tgccaacctt ttcttttata aaactattac atttaaaata tggacgtctg 1260
aaaaataaaa tattcatcat ttttatgaaa aaaaaaaaaa aa

```

&lt;210&gt; 45

&lt;211&gt; 1771

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 130117CB1

&lt;400&gt; 45

```

tgacttacag ctcttataaa ctagtggcaa tttctgaacc cagccggctc catctcagct 60
tctggtttct aagtccatgt gccaaaggct gccaggaagg agacgccttc ctgagtcctg 120
gatctttctt ccttctggaa atctttgact gtgggtagtt atttatttct gaataagagc 180
gtccaogcat catggacctc gcgggactgc tgaagtctca gttcctgtgc cacctgggtc 240
tctgtcagct ctttattgac tcagggctaa tcatcaacac cattcagctc ttcactctcc 300
tctctggcc cattaaacag cagctcttcc ggaagatcaa ctgcagactg tctattgca 360
tctcaagcca gctgggtgat ctgctggagt ggtggctcggg cacggaatgc accatcttca 420
cggaccgcgc cgctacctc aagtatggga aggaaaatgc catcgtgggt ctcaaccaca 480
agtttgaaat tgactttctg tgtggctgga gcctgtccga acgctttggg ctgttagggg 540
gtccaaggct cctggccaag aaagagctgg cctatgtccc aattatcggc tggatgtgg 600
acttcaccga gatggctctc tgttcgcgca agtgggagca ggatcgcaag acggttgcca 660
ccagtttgca gcacctccgg gactaccccg agaagtattt ttctctgatt cactgtgagg 720
gcacacgggt cacggagaag aagcatgaga tcagcatgca ggtggcccg gccaaggggc 780
tgccctgcct caagcatcac ctgttgccac gaaccaaggg cttcgccatc accgtgagga 840

```

```

gcttgagaaa tgtagtttca gctgtatatg actgtacact caatttcaga aataatgaaa 900
atccaacact gctgggagtc ctaaaccggaa agaaatacca tgcagatttg tatgttagga 960
ggatcccact ggaagacatc cctgaagacg atgacgagtg ctgggctgg ctgcacaagc 1020
tctaccagga gaaggatgcc ttccaggagg agtactacag gacgggcacc ttcccagaga 1080
cgcccatggg gcccccccg cgcccttgga ccctcgtgaa ctggctgttt tgggctcgc 1140
tggtgctcta ccctttcttc cagttcctgg tcagcatgat caggagcggg tcttccctga 1200
cgctggccag cttcctcttc gtcttctttg tggcctctgt gggagtcca tggatgattg 1260
gtgtgacgga aattgacaag ggctctgcct acggcaactc tgacagcaag cagaaactga 1320
atgactgact cagggaggtg tcaccatccg aagggaacct tggggaactg gtggcctctg 1380
catatctctc ttagtgggac acggtgacaa aggcctgggtg agccctgct gggcacggcg 1440
gaagtcacga cctctccagc cagggagttc ggtctcaagg ccggatgggg aggaagatgt 1500
tttgtaatct ttttttccc atgtgcttta gtgggctttg gtttctttt tgtgcagtg 1560
tgtgtgagaa tggctgtgtg gtgagtgtga actttgttct gtgatcatag aaagggtatt 1620
ttaggctgca ggggagggca gggctgggga ccgaagggga caagttccc tttcatcctt 1680
tggtgctgag tttctgtaa cccttgggtg ccagagataa agtgaaaagt gctttaggtg 1740
agatgactaa attatgcctc caagaaaaaa a 1771

```

&lt;210&gt; 46

&lt;211&gt; 1755

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 267495CB1

&lt;400&gt; 46

```

caatccgaag gacaggaagg ccattcttaga ttgataagcc atcttatgtt cccagtatct 60
gtctgaaata attatatcct gagagctctc gtagtgtgaa gctttcttct acatgtaaaa 120
tttcacaggg aacctatggtt ttcaaacaaa ggcataagaa cccaggctta ggaagcgggt 180
gtttaagcca tgggttagag ggctcgtgag cagcccgat cctttccctg aggctatgaa 240
acgggtggctc cacggactcc tctcgccga ctgaccttc gccgccctgc cccagcagcc 300
ggcgggtttc ttcagtggag cgggctctg gtctccgcag ccagtagcc cgctagccc 360
gccccctccc gctatcgta aagtgcagta aaggcaccag cattttgagg caccgtatgt 420
aggccggcgg cgtgggcctt ttctctgcac ggagccggcg cttttgcagt tgettctgag 480
gaaagggtgt agttaagaat ttgtaaaggc cagagaacta cctacgattc tctcagcgtt 540
ctctcttctc ctcaagtttg aaatgcttta tctcatcggt ttgggctgg gagatgcaa 600
ggacatcaca gtcaagggcc tggaaagtgt tagacgtgc agtcgagtg atctggaagc 660
ctacacctca gtccaaactg tagggaagga agccttgga gagttttatg gaagaaaatt 720
ggttggtgct gatagagaag aagtggaaac agaagcagat aatattttta aggatgctga 780
tatcagtgat gttgcattcc ttgtggttgg tgatccattt ggggccacaa cacacagtga 840
tctgtttcta agagcaacaa agctgggaat tccttataga gttattcaca atgcctccat 900
aatgaatgct gtaggtgctt gtggtttaca gttatataag tttggagaga cagtttctat 960
tgttttttgg acagacactt ggagaccaga aagcttcttt gacaaagtga agaagaacag 1020
acaaaatggc atgcacacat tatgtttact agacatcaa gtaaaggagc agtctttgga 1080
aaatctaata aagggaagga agatctatga acctccacgg tatatgagtg taaaccaagc 1140
agcccagcag cttctggaga ttgttcaaaa tcaaagaata cgaggagaag aaccagcagt 1200
taccgaggag acactttgtg ttggcttagc cagggttgga gccgacgacc agaaaattgc 1260
agcaggcact ttaaggcaaa tgtgcactgt ggacttgga gaaccattgc attccttgat 1320
catcacagga ggcagcatac atccaatgga ttggagatg ctaagtctgt ttccatacc 1380
agaaaatagc tcagaatctc aaagcatcaa tggactttga acatagatat ttaccattgt 1440
ctgatgtaaa tttcagccat atatggattg atatggtttg gatgtatccc cacccaagtc 1500
tcactctgaa ttttaatcct cataattccc aggtgttgtg gtaggtaatt gaatcatggg 1560
ggcagtttcc ctcatgctat tctcatgata gtgagcttcc atgagatctg atggttttat 1620
aagtgcctgg catttcccct actggctctc attctcactc ttgccgcct gtgaagaggt 1680
gccttcacc gtgattgtta agtttctga ggccttccca gccatgtgga actgtgagtc 1740
gaaaattaac cctct 1755

```

&lt;210&gt; 47

&lt;211&gt; 1811

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 410533CB1

&lt;400&gt; 47

```
gcggcgccag ggccatgcta gccttgccgc tggcgccgcg ctcgtggggg gccctgcgcg 60
gcgcgccttg ggctccggga acgcggcgag gtaagcgacg cgctgctgg gccctgctgc 120
cgcccgctgc ctgctgcttg ggctgccttg ccgaacgctg gaggtgcgt ccggccgctc 180
ttggcttgcg gctgcccggg atcgcccgag ggaaccactg ttcgggcgcg gggaaggcgg 240
ctcccaggcc agcggccgga gcgggcgcgc ctgccgaagc cccgggcggc cagtggggcc 300
cggcgagcac cccagccctg tatgaaaacc catggacaat cccgaatatg ttgtcaatga 360
cgagaattgg cttggcccca gttctgggct atttgattat tgaagaagat ttaatatg 420
cactaggagt .ttttgcttta gctggactaa cagatttggt ggatggattt attgctcgaa 480
actgggccaa tcaaagatca gctttgggaa gtgctcttga tccacttgct gataaaatac 540
ttatcagtat cttatatgtt agcttgacct atgcagatct tattccagtt ccacttactt 600
acatgatcat ttcgagagat gtaatgttga ttgctgctgt tttttatgtc agataccgaa 660
ctctccaac accacgaaca cttgccaaat atttcaatcc ttgctatgcc actgctaggt 720
taaaaccaac attcatcagc aagggtgaata cagcagtcga gttaatcttg gtggcagctt 780
ctttggcagc tccagttttc aactatgctg acagcattta tcttcagata ctatgggtgtt 840
ttacagcttt caccacagct gcatcagctt atagttacta tcattatggc cggaagactg 900
ttcaggtgat aaaagactga tgaaagtcac ccctcactgt tagtaaggaa gcagtataca 960
tcaatgggaa cagggcccat ggaaatgtac aggagtttcc ctattttggt gttcagcttg 1020
aaaaaggact tgtcagaatc aactgtgtca tcaaaattta agtaatgtgc attgaaaata 1080
aggttgatca tgggaatatg cagaatttcc aatgtatttt taaatacaaa taaaattgta 1140
atttagaatt tttaatctta ggtttcttga ttaatttata agagatcaat tattgtcagt 1200
cttttttgta tgttttttaa aaacatagtc cagagcatgg gcagaattga cacctctctt 1260
ttaagtgaat tttggattgc tcacaaagca ctaggaaatg tcatgggggtt caaatatata 1320
tcctacacaa ctgggcaata catttttggt tgatttttag gtctgtgtat acattaacag 1380
ttcatgtaat taatacttga tcatttggga taatgaaagt gaagttagtt gtagatgaag 1440
taaagttata aaagagatta aaaatgcggt aactttttta gataataatc atacagaagg 1500
tatgaagttc attttcggta gtcttccaac ctctcagggt cctaataatt tatgtttgag 1560
gataacaggt aacaaagata gtcgagatag gagaacgtgt ctattagtct ttgcatctaa 1620
aaggcagtg tttacgttcc tgccttccac tgctgttctg acatagcaat gtttgtttga 1680
tattggaacc ggtattcata ttttatgtaa ataatatcaa gctgtatatt tttcaaaggt 1740
tttttaaaact ttggagactc tttcttttgt taagcagtta aaggataaaa agagctggaa 1800
aaaaaaaaa a 1811
```

&lt;210&gt; 48

&lt;211&gt; 1003

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 852708CB1

&lt;400&gt; 48

```
tcccacgcgt ccgcggaacgc gtgggcggac gcgtgggcgg acgcgtgggg cggtggtcca 60
ggctccttga aaggaccgtc caccctctcg cgctggcggt gtggacgcgg aactcagcgg 120
agaaacgcga ttgagaaatg gaaaagaaaa tgaaataaat cagcagttat gaggcagagc 180
ctaagagaac tatggcaaca tcaggtgact gtcccagaag tgaatcgag ggagaagagc 240
ctgctgagtg cagtggggcg ggtctcctgc aggagggagt acagccagag gagtttgtgg 300
ccatcgcgga ctacgctgcc accgatgaga ccagctcag ttttttgaga ggagaaaaaa 360
ttcttatect gagacaaacc actgcagatt ggtggtgggg tgagcgtgcg ggctgctgtg 420
ggtacattcc ggcaaacccat gtggggaagc acgtggatga gtacgacccc gaggacacgt 480
ggcaggatga agagtacttc ggcagctatg gaactctgaa actccacttg gagatggttg 540
cagaccagcc acgaacaact aaataccaca gtgtcatcct gcagaataaa gaatccctga 600
cggataaaat catcctggac gtgggctgtg ggactgggat catcagtctc ttctgtgcac 660
actatgcgcg gcctagagcg gtgtacgcgg tggaggccag tgagatggca cagcacacgg 720
ggcagctggg cctgcagaac ggctttgctg acatcatcac cgtgtaccag cagaaggttg 780
aggatgtggg gctgcccagc aagggtggag tgctgggtgtc tgagtggatg gggacctgcc 840
tgctgggtgag ggccggcgctg cgggcagctg ggggcccggg ctggggggct tctgagcagc 900
ggctcggtcg ggtccaacctc aggatctcaa gggtcggtgc tgattcattt tgatgttttc 960
cctaattgtga ggtctaatta atttcttgtt tggaaaaaaa aaa 1003
```

&lt;210&gt; 49

&lt;211&gt; 1687

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 972944CB1

&lt;400&gt; 49

```
gctgccacgg taacgcgcgc gggcaggtgt cgcaccatga gcgtccgggt cgcacgggta 60
gcgtgggtca ggggcttggg cgcagctac cgcgcggcg cctcgagctt cccggtgcct 120
ccgcggggcg cccaggtgt agcggagctg ctgcgagatg cgaccggggc ggaggaggag 180
gcgcctggg cggcgacgga gcggcgaatg cggggccagt gctccgtgct gctcttcccg 240
ggccagggca gccaggtgt gggcatgggc cgcggtctgc tcaactaccg gcgcgtccgc 300
gaactctacg ccgcgcgccg ccgcgtgctg ggctacgacc tgcagggaact gagcctgcac 360
gggcccaggg agaccctgga ccgcaccgtg cactgtcage ccgcgatctt cgtggcatcg 420
ctggccgctg tcgagaaact acatcacctg cagccctcgg tgattgagaa ctgtgttgc 480
gctgctggat tcagtgtggg agagtttgca gccctagtgt ttgccggagc catggaattt 540
gctgaagggt tgtatgcagt gaaaatccga gctgaggcca tgcaaggagc ttcagaagct 600
gtccccagtg ggatgctgtc tgtcctcggc cagcctcagt ccaagtcca cttcgctgt 660
ttggaagccc gggaacactg caagtcttta ggcatagaga acccgtatg tgaagtgtcc 720
aactacctct tccagaagct cagggtgatt tcaggacacc aagaggctct acggtttctc 780
cagaagaatt cctctaagtt tcatctcaga cgcaccagga tgttgccggt tagtggcgca 840
ttccacaccc gcctcatgga gccagccgtg gagccctga cgcaagcttt aaaggcagtc 900
gacattaaga agcctctggt ttctgtctac tccaacgtcc acgcgcatag atacaggcat 960
cccgggcaca cccacaagct gctggccagg cagctggctc cccagtgaa gtgggagcag 1020
acgatgcacg ccatatacga aaggaaaaag ggcagggggg tcccccaaac tttcgaagta 1080
ggccctggca ggcagctggg agccatcctg aagagctgta acatgcaggc ctggaagtc 1140
tacagcgccg tggatgtgct gcagaccctc gaacatgtgg acctggaccc tcaggagccc 1200
ccgagatgac tgcagggggc tcaaatgcga tgacccctc tgcctcctg agggagggt 1260
gtaggctgtg cctgtcgccc cctaccttcc taatggctcc tectctgagg agtgaagggt 1320
atgtgttgac aacgtgcttt gaaggccaca taaaagccc taaaatgag tatttcttta 1380
catgaccagc tccatttctc ccccttgaa aaacgtttgg acgttgggaa gaatgatgcc 1440
acagggtggt tgtgtgggga agctggaatc ctggcccgcc ctctgccagc cctgcacct 1500
aggcaggtgt agaacatgct ggtgtggcct caatttcagc gctgcagcct gtctgtctct 1560
gtgagtgctg agaacatgct ggtgtggcct caatttcagc cagtgcagg ggagccctc 1620
caccacccc agcctgcccc ccagggtcc tttgtcaca atctgaaagg ttttagaaaa 1680
aaaaaaa
```

&lt;210&gt; 50

&lt;211&gt; 1239

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 997730CB1

&lt;400&gt; 50

```
ctcagaggc cgtgggtacg accggaagcc gcagcatgct gatcccatTT tcaatgaaga 60
attgcttcca gttactttgt aactgccagg tcccagcagc tggctttaa aaaacagtaa 120
aaaatgggct cattttacag tcaatttcca atgatgtcta tcaaaatctg gctgtggaag 180
actggatcca tgaccatag aatctagaag gcaaaccaat tctattcttt tggcagaatt 240
ctccctctgt tgtaattggt aggcatacaa atccttgga ggaatgtaac ctgaattcaa 300
tgagagaaga aggtataaaa ctggctcgga gaagaagtgg aggaggaaca gtctaccatg 360
atatgggtaa tatcaatttg actttcttta caacaaaaaa aaagtatgat agaattgaaa 420
atctgaaatt aattgtgaga gctctgaatg ctgtccaacc ccagctggat gtgcaggcta 480
ccaaaagatt tgacctttta cttgatggac agtttaaaat ctcaggaaca gcttctaaga 540
tcggccggac tactgcctat caccattgca ctttattatg tagtactgat gggacgttcc 600
tgtcttcttt gctaaagagc ccttaccagg ggatcaggag caatgccact gctagcatac 660
cttctctagt gaaaaatctt ttggaaaagg atcccactct gacctgtgaa gtactaatga 720
atgctgttgc tacagagtat gctgcctatc atcaaatgta taatcacatt cacctaataa 780
acccaacgga tgagacactg tttcttgga taaatagcaa agccaaagaa ctgcaactt 840
gggagtggat atatggcaaa actccaaagt ttagtataaa tacttcttt catgtgttat 900
atgaacagtc acacttgga attaaagtat tcatagacat aaagaatgga agaattgaaa 960
tttgtaatat tgaagcact gatcattggt tgccattgga aatacgtgac aaattaaatt 1020
caagtcttat tggcagtaag ttttgcccaa ctgaaactac catgctaaca aatatattac 1080
ttagaacatg tccacaagac cacaactaa acagtaaatg gaattattct tgtgaaaaaa 1140
ttaagggaat aatgtgattc caagtaaatg tcttaataca gtttcaatta gaaaataaaa 1200
tgtctcatal ttgcacattg tatgtcaaaa aaaaaaaaa 1239
```

&lt;210&gt; 51

<211> 1594  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1285944CB1

<400> 51  
gtcggagccc taaccagggg tatctctgag cctgggtggga tccccgggag gtcacatcac 60  
tttccgatca cttcaaaagt gttaaaaaact aatattttata tgacagaaga aaaagatgtc 120  
attccgtaaa gtaaacatca tcatcttggg cctggctgtt gctctcttct tactggtttt 180  
gcaccataac ttcttcagct tgagcagttt gttaaggaat gaggttacag attcaggaat 240  
tgtagggcct caacctatag actttgtccc aaatgctctc cgacatgcag tagatgggag 300  
acaagaggag attcctgtgg tcatcgctgc atctgaagac aggcctgggg gggccattgc 360  
agctataaac agcattcagc acaacactcg ctccaatgtg attttctaca ttgttactct 420  
caacaataca gcagaccatc tccggctcctg gctcaacagt gattccctga aaagcatcag 480  
atacaaaatt gtcaattttg accctaaact tttggaagga aaagtaaagg aggatcctga 540  
ccagggggaa tccatgaaac ctttaacctt tgcaagggtc tacttgccaa ttctgggttc 600  
cagcgcaaag aaggccatat acatggatga tgatgtaatt gtgcaagggt atattcttgc 660  
cctttacaat acagcatcga agccaggaca tgcaagctgca ttttcagaag attgtgattc 720  
agcctctact aaagtgtgca tccgtgggagc aggaaaccag tacaattaca ttggctatct 780  
tgactataaa aaggaaagaa ttctgtaagct ttccatgaaa gccagcactt gctcatttaa 840  
tcctggagtt tttgttgcaa acctgacgga atggaaaacga cagaatataa ctaaccaact 900  
ggaaaaatgg atgaaactca atgtagaaga gggactgtat agcagaacct tggctggtag 960  
catcacaaca cctcctctgc ttatcgtatt ttatcaacag cactctacca tcgatcctat 1020  
gtggaatgtc cgccaccttg gttccagtgc tggaaaacga tattcacctc agtttgtaaa 1080  
ggctgccaag ttactccatt ggaatggaca tttgaagcca tggggaagga ctgcttcata 1140  
tactgatgtt tgggaaaaat ggtatattcc agaccaaca ggcaaatcca acctaatccg 1200  
aagatatacc gagatctcaa acataaagtg aaacagaatt tgaactgtaa gcaagcattt 1260  
ctcaggaagt cctggaagat agcatgcgtg ggaagtaaca gttgctaggc ttcaatgcct 1320  
atcggtagca agccatggaa aaagatgtgt cagctaggta aagatgacaa actgcctgt 1380  
ctggcagtc gcttcccaga cagactatag actataaata tgtctccatc tgccttacca 1440  
agtgttttct tactacaatg ctgaatgact ggaaagaaga actgatatgg ctagtccagc 1500  
tagctggtag agataattca aaactgctgt tggttttaat tttgtaacct gtggcctgat 1560  
ctgtaaataa aacttacatt tttcaaaaaa aaaa 1594

<210> 52  
<211> 700  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1293207CB1

<400> 52  
cccggcctgc ggtgggcagc agctcaggtt ctccaaatca ttgcgtagtt ccgaataccc 60  
tcggccacac ctggccttct ccatgctcgg aataacttcc tgcagcgacc aacaggctaa 120  
agagggggaa gggatccagc accggctcct cctccggcaa ccacgggtggg agcggcgag 180  
gaaatggaca taaacccggg tgtgaaaagc cagggaatga agcccgcggg agcggggaat 240  
ctgggattca gaactctgag acgtctcctg ggatgtttaa ctttgacact ttctggaaga 300  
attttaaatc caagctgggt ttcatcaact gggatgccat aaacaagaac cagggtcccgc 360  
ccccagcac ccgagccctc ctctacttca gccgactctg ggaggatttc aaacagaaca 420  
ctccttttct caactggaaa gcaattattg aggggtcgga cggtcatca ctgcagaaac 480  
gtgcaggcag agccgatcag aactacaatt acaaccagca tgcgtatccc actgcctatg 540  
gtgggaagta ctcagtcaag acccctgcaa aggggggag ctcaccttct tcctcggtt 600  
cccgggtgca acctggcctg ctgcagtggg tgaagttttg gtaggtagt gtcagagtga 660  
gccgaccag gccacatcct ggcagtggag gcacagtcac 700

<210> 53  
<211> 536  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature



&lt;223&gt; Incyte ID No: 1308125CB1

&lt;400&gt; 53

gttttaggcg	agttggagat	gtctgtgaga	ggcttgaaca	gtgatgtcca	gttcaaggat	60
ggagggcaaa	gccaagtaca	tcctccctac	cgaacaata	tatgttgggg	aatgaagga	120
tggcatgttt	cacggcgagg	gaaccctgta	cttccccagc	ggaagccaat	acgacgccat	180
ttgggaaaac	ggattggcca	taaaggatat	gctcaactca	ccaatatgga	cccacctaga	240
aaaaatcccc	agggctatta	cgattgtgga	gacggcttct	ataaccagt	cacgagggta	300
gtcaaggact	ataggaaccg	ctttctaaga	aacgcagatg	atgacgagca	tgagtggatc	360
acccgtacct	gtcgaaaggg	ctaggatgag	atcgtgggtc	acaggcccga	gccgtgaact	420
ctgtggctgc	ctccaccaga	ggtttccatc	tgccctacta	gcattggctg	ccctggggga	480
cgggctgtag	ttctagaacc	tgattttaac	tcaggaataa	agactttctg	cgggtca	536

&lt;210&gt; 54

&lt;211&gt; 1130

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1439670CB1

&lt;400&gt; 54

agaggcgagg	tgaggcggtt	ggcgctgcc	cgtctggg	gcggttccca	actgtggcgc	60
gggcggtgga	ggaggaggtg	gggctggcgc	tgaagccgga	tccggatccg	gtgctgtgca	120
cactggtggg	ggagagtccg	acgcgcctgg	ctaggagcgc	cgaccgcagg	gcctctacgg	180
accttactag	aaaaatgaaa	cctgatgaaa	ctcctatggt	tgaccaagt	ctactcaaag	240
aagtggactg	gagtcagaat	acagctacat	ttctctcagc	catttcccca	acacatcctg	300
gagaaggctt	ggttttgagg	cctctttgta	ctgctgactt	aaatagaggt	ttttttaagg	360
tattgggtca	gctaacagag	actggagttg	tcagccctga	acaatttatg	aaatcttttg	420
agcatatgaa	gaaatctggg	gattattatg	ttacagttgt	agaagatgtg	actctaggac	480
agattgttgc	tacggcaact	ctgattatag	aacataaatt	catccattcc	tgtgctaaga	540
gaggaagagt	agaagatgtt	gttggttagtg	atgaatgcag	aggaaagcag	cttggcaaat	600
tgttattatc	aacccttact	ttgctaagca	agaaaactgaa	ctgttacaag	attacccttg	660
aatgtctacc	acaaaatgtt	ggtttctata	aaaagtgttg	atatactgta	tctgaagaaa	720
actacatgtg	tcggagggtt	ctaaagtaaa	aatcttgtaa	gaaaattgtc	aaaggggcta	780
atgctacaag	gctacactct	tcctagagtt	gaaatatatt	gttgctgcag	ccgagtgcac	840
tccataaata	ctggactgaa	aaaacattgt	aatactacaa	gtataatgac	atttagaaga	900
ttactttggg	ctgggtgggac	atgctgtgaa	tttagattac	aatgaatat	tataacaggg	960
gatgattttt	aaccaaacgg	aatatatatt	taacttggat	cttttcttgc	attgtatttt	1020
ttctaaaagg	tttggcatcc	ctatcttggg	aaggtcagga	ggtatggggt	aaataaaggg	1080
agttaataat	ggtctgggta	atccgggtggt	ttggcctcat	tttataaaaa		1130

&lt;210&gt; 55

&lt;211&gt; 930

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1444281CB1

&lt;400&gt; 55

ctccgagagc	gggcggggct	cagttcagct	gctgtccaga	cccggatcgg	caacagtgcc	60
gectccagac	gttctcctgc	cgctcgcccg	cccgtccag	cgccccagc	cctcccgcca	120
gggcgccccg	ggacggaagg	atccaccagt	ctgtcggcgc	cgcgcgttct	cgtggtcgcc	180
gtcgccgtcg	tcgtggtggg	agtctccgcc	gtcgccctgg	ccatggccaa	ttacatccac	240
gtccctcccg	gtcctccgga	gggtgcccaag	ctgaacgtca	ccgttcagga	tcaggaggag	300
catcgctgc	gggagggggc	cctgagccct	ctgcaacacc	tgcgccctca	ctgggacccc	360
caggaggtga	ccctgcagct	cttcacagat	ggaatcacaa	ataaacttat	tggctgttac	420
gtgggaaaca	ccatggagga	tgtagtccct	gtgagaattt	atggcaataa	gactgagtta	480
ttagtgcgac	gagatgagga	agtaaagagt	tttcgagtgt	tgcaggctca	tgggtgtgca	540
ccacaactct	actgtacctt	caataatgga	ctatgctatg	aatttatata	aggagaagca	600
ctggatccaa	agcatgtctg	caaccagcc	attttcagtt	tatcatcggt	gactctttgc	660
aaaggaaaaa	ctacaagatg	ttttggatta	acgggctgca	gagggtaag	gcttctgctt	720
agttttttct	agtttagtgt	tgagttgaag	tgttatcat	ttcttactgc	cgattgctta	780
ttacttaatt	gttcatgttt	tcagagttct	tgcctattaa	aattttatta	atttgggtaa	840

caattcttttg ttctttacgt tgttcattat ttaatttaat atattatggt ttctataaag 900  
aataatgtca tatatgcttg atcatttttg 930

<210> 56  
<211> 1687  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1450140CB1

<400> 56  
gtccgcgggg agctgcccct gtgcccgtgg ctctctgcgc atgcccggac cccggcggct 60  
ccggagtcgc ggaggtctctg gctgctccag tcgcgggtcc gggccccgcc tctccactcg 120  
gcgggtggcgt ccgtgttgtt attgtggtgg tagcggcgctc tggctgctgc ggaccccgcc 180  
gagtcctagc gectggcctg cgcgcgcctg cccgcgccac aggattaatt tatttttgga 240  
aatcaagtgc aatattggaa gccatttcta ctaattttat ttcactttta aatttatgat 300  
ttgatttgaa tactatcatg ggaggagctg tgagtgtctg ggaagataat gatgacttaa 360  
ttgataattt aaaagaagct cagtatattc gtactgaaag agtggagcaa gccttcagag 420  
cgattgatcg tggagattac tatttggaag gctacagaga caatgcttac aaagacttag 480  
cctggaagca tggaaacatc cacttgtcag cacttgcac ttattctgaa gttatggaag 540  
cattgaaact tcaaccagga ttgtcttttc ttaacctggg aagtgggaacc ggatatttaa 600  
gtacaatggt gggcttaatt ttaggtcctt ttggaataaa tcatgggatt gagcttcatt 660  
cagatgtggt ggaatatgcc aaggaaaaac tggagagctt catcaaaaat agtgatagct 720  
ttgataaatt ttgagttctg gaacctgcat ttgttgttgg taattgcctc cagatagctt 780  
ctgacagtca tcagtatgat cgaatttatt gtggagctgg agtacagaaa gacctgaaa 840  
actacatgaa aatattacta aaagttggag gcatattagt catgcctata gaggatcagt 900  
taacacagat tatcggaact ggacagaaca ctggggaaag taaaaatata cttgctgttt 960  
catttgctcc acttgtgcaa ccaagtaaga atgataatgg caaaccagat tctgtgggac 1020  
tccctccctg tgctgtcagg aatctacagg acttggctcg tatttacatt cgacgcacac 1080  
ttagaaaatt cataaatgat gagatgcagg ccaaggggat tcctcaaagg gctccacca 1140  
aaaggaaaaa aaagagagtt aaacagagaa ttaacactta cgtatttgtg ggtaatcagc 1200  
ttattcctca gcctctagac agtgaagagg atgaaaaaat ggaagaggat atcaaaagag 1260  
aggaggaaaa agatcacaat gaagcaatga agccagagga gccacctcaa aatttactga 1320  
gagaaaaaat catgaagctg cccctccctg aatctttaaa agcttacttg acatatttta 1380  
gagacaaata acttagatca agaagaaaaa tgcctactga taattccttt agtcttgaaa 1440  
atgtagcatt tgttaggagt taaaagagag aattatttct ttcattcagag caaattatag 1500  
tggaaaaaaa aatcacttgt ttctgtcagt aacacaaatg atgtattcag tgaataaaag 1560  
aatccctttt ataaaatcta tttttcttta aatcttggaa aaatgttgtt ttagctcaga 1620  
gtgatttcaa agtggaatgc aacagtagtc aagacttgtg tactataaat ctttttctga 1680  
ttcetta 1687

<210> 57  
<211> 1262  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1604828CB1

<400> 57  
ggacgcgcgc cctcctctcg ggcggagcgg cgcgggtggct gatcagagcg cgtagggtt 60  
gcgcgggtgc ggcgctctgg gcggtcctg ctccagcccag ctaccgcgc gccggccctc 120  
gggcgccttg ttctgcggat caggagaaaa taatgaatgt cagaggaaaa gtaattctgt 180  
caatgctggt tgtctcaact gtgatcattg tgttttggga atttatcaac agcacagaag 240  
actctttctt gtggatatat cactcaaaaa acccagaagt tgatgacagc agtgtcaga 300  
agggctggtg gtttctgagc tgggttaaca atgggatcca caattatcaa caaggggaag 360  
aagacataga caaagaaaaa ggaagagagg agaccaaagg aaggaaaatg acacaacaga 420  
gcttcggcta tgggactggt ttaatccaaa ctgaaggaa tccgaataac taaactggac 480  
tctgttttct tgactcagtc cttctagaag acctggactg agagatcatg cggttaagga 540  
gtgtgtaaca ggcggaccac ctgttgggac tgcgagattc tcaaggggaa ggactgggtc 600  
tcatttctcc catctcagcg cttagcagga tgacctggta tagagcaggg aactgggaaa 660  
tgtgggtcag gggatcagac actccagttg ggtcttttat ataaattaaa tggcaaaaag 720  
ctccataccc ttctccttct ttctaccct ccactttatc tgcaaaatgg gaatgatgat 780  
aacaccact tcatagaatg gtcatgaaga tcaaatgaga gaataaaagt caagcactta 840

```

gcctctggtg cacaataagt attaaataag tatacctatt cctccttttc ctttttttaa 900
aataatatta ccaaagtcc agcttatata cattttacaag acttagctag tgggctatgt 960
tagagctact aaaagatctt tgacaagcta aaactaagat gcaatgaatg aggtgtaacg 1020
aacaagagag ttttaagttc agaaatggtt acagaagtat aagacagctg tgtgggtgtt 1080
ttttggtttt tgggtttctg tttacaatct cgtcattcaa caaagatggg agttttatag 1140
aactaaaagc accatgtaag ctactaaaaa caacaacaaa aaaggctcat cattttctag 1200
tctgaattga caaaaatgcc aatgcaata aaaaatgatta ctttttattt taaaaaaaaa 1260
aa 1262

```

<210> 58  
 <211> 1330  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1644023CB1

```

<400> 58
gcgaagcggg cgtgcgccag tgcagtgcgc cgggcgcgc gcgactcccg atagggaagg 60
gccggagctg cgcagggggc cgcggaagag gccaaagctag ccggaacccc gccccgcccc 120
gccccgcccc aacgggggct ggaacccggc gccgagagta gagaaaaggg gcctctgggtg 180
accgccccct cctggcatcc ctctaaccct ggaggagcgt ggggaaaggg gctgtggggc 240
tctcggggag cgagctgcgc gtagcggcgc actgggtaca ggcgcgcgct tggctgtcgc 300
ctctgccgct gtgtttggga ggactcgac tggcgccagg aaatattagg aagctgtgat 360
tttcaaagct aattatgaaa acatttatca ttggaatcag tgggtgtgaca aacagtggca 420
aaacaacact ggctaagaat ttgcagaaac acctcccaaa ttgcagtgtc atatctcagg 480
atgattttct caagccagag tctgagatag agacagataa aaatggattt ttgcagtacg 540
atgtgcttga agcacttaac atggaaaaaa tgatgtcagc catttcctgc tggatggaaa 600
gcgcaagaca ctctgtggta tcaacagacc aggaaagtgc tgaggaaatt cccattttta 660
tcatcgaagg ttttcttctt ttttaattata agccccttga cactatatgg aatagaagct 720
atttcttgac gattccatat gaagaatgta aaaggaggag gagtacaagg gtctatcagg 780
ctccagactc tccgggatac tttgatggcc atgtgtggcc catgtatcta aagtacagac 840
aagaaatgca ggacatcaca tgggaagtgt tgtacctgga tggaaacaaa tctgaagagg 900
acctcttttt gcaagtatat gaagatctaa tacaagaact agcaaagcaa aagtgtttgc 960
aagtgcagac ataaagacgg aacacaacaa atccttctct aagtgaatta gaaacttcca 1020
aggagtaatt taagaacctt caccaagata caatgtatac tgtggtacaa tgacagccat 1080
tgtttcatat gtttgatttt tattgcacat ggttttccca acatgtggaa caataaatat 1140
ccatgccaat ggacaggact gtaccttagc aagttgtctc ctctccaggg agcgcataga 1200
tacagcagag ctacacagtga gtcagaaagt ctccacttcc tgaacatagc tctataacaa 1260
tgattgtctaa accttttctaa ctggagctca gagtaagaaa taaagattac atcacaatcc 1320
aaaaaaaaa 1330

```

<210> 59  
 <211> 1110  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1723402CB1

```

<400> 59
taagctgaat cgactgctgc caaacatcta ttaggcaaaa ttggcctctt gcccatgatt 60
tgactttcca gcacagccag ttcttttctt cctctgcagc tgattggctc tggagtgtgg 120
ccagaagcct ctctcctgca attaaaggag tcgggtctct aactgttgat ctgttttttt 180
cccttctgag caatggagct taccatcttt atcctgagac tggccattta catcctgaca 240
tttcccttgt acctgctgaa ctttctgggc ttgtggagct ggatatgcaa aaaatgggtc 300
ccctacttct tggtaggtt cactgtgata tacaacgaac agatggcaag caagaagcgg 360
gagctcttca gtaacctgca ggagtttgcg ggcccctccg ggaaactctc cctgtgggaa 420
gtgggctgtg gcacgggggc caacttcaag ttctaccac ctgggtgcag ggtgacctgt 480
attgacccca accccaactt tgagaagtgt ttgatacaaga gcattgcaga gaaccgacac 540
ctgcagtttg agcgccttct ggtagctgcc ggggagaaca tgcaccaggt ggctgatggc 600
tctgtggatg tgggtgtctg caccctgggt ctgtgctctg tgaagaacca ggagcggatt 660
ctccgcgagg tgtgcagagt gctgagaccg ggaggggctt tctatttcat ggagcatgtg 720
gcagctgagt gttcgacttg gaattacttc tggcaacaag tcctggatcc tgccctggac 780
cttctgtttg atgggtgcaa cctgaccaga gagagctgga aggccctgga gggggccagc 840

```

```

ttctctaagc tgaagctgca gcacatccag gccccactgt cctgggagtt ggtgcgccct 900
catatctatg gatatgctgt gaaatagtgt gagctggcag ttaagagctg aatggctcaa 960
agaattttaa gcttcagttt tacattttaa atgctaagtg ggagaagaga aacctttttt 1020
ttggggggcg gtttttttgg tttgttgttg gttttttttt tttttttggc ctgggtgaca 1080
agagcaagac tccgtctcaa aaaaaaaaaa 1110

```

<210> 60  
 <211> 1153  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1740585CB1

```

<400> 60
ggtagagcgg agacgacgct cccagactcc tccggtctcc ccgggcagca tgaagaccgc 60
cgagaacatc agaggaaccg gcagcgacgg gccgcggaaa cgaggcctct gcgtcctctg 120
tggcctcccc gcggcaggaa aatcgacttt cgcgcgcgcc ctcgccacc ggctgcagca 180
ggagcagggt tgggcatcg gtgttgctgc gtatgatgac gtcatgccg acgcgtttct 240
cgccggggca agagcgcgac cggcgccatc ccaatggaaa ttgcttcgac aggaactgtt 300
gaagtacctg gaatacttct tgatggctgt cattaatggg tgtcagatgt ctgtcccacc 360
caacaggact gaagccatgt gggaagattt tataacctgc ttaaaggatc aagatctgat 420
attttctgca gcatctgagg cccagtcttg ctacctcta acaaaaactg ctgtttctag 480
acctttgttt ttggttttgg atgacaattt ttattatcag agtatgagat atgaagtcta 540
ccagctggct cggaaatatt cattgggctt ttgccagctc tttttagatt gtctcttga 600
gacctgttta cagaggaatg gccagagacc acaggcactg cctcctgaga ccatccacct 660
gatgcgaaga aagctagaaa agcccaaccc tgagaaaaat gcttgggaac acaacagcct 720
cacaattccg agtccagcat gtgcttcgga gccagcctg gaagtactg atttattgct 780
cactgctttg gaaaatccag taaaatatgc tgaggacaat atggaacaaa aggacacaga 840
cagaattatt tgttcaacta acattcttca taaaactgat cagacactcc gaaggattgt 900
atctcagaca atgaaggaag caaaagatga acaagtgtt cctcacaact tgaagcttct 960
agcagaagaa cttaacaagc tcaaagcaga gtttttgaa gacctaaaac aaggaaaaca 1020
aaaaatatct tgctttcagc aaaccattga cataccagat gtcatttctt tttttcatta 1080
tgagaaagat aatattgtac agaagtattt ttcaaagcag cattaataatt tctgaactgc 1140
caaaaaaaaa aaa 1153

```

<210> 61  
 <211> 1955  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1810925CB1

```

<400> 61
gggcaggcta agagatcttc ttttaattca gcctgcttaa gacgggaact gataactgta 60
gtgtatcttc tgcctttttt ctatctatt ggaggaagct cagatggtgt cacaagaagg 120
atctgaagtg gagcttctag tatccccagg agcgcgaagt gaacacggaa ggtacctgca 180
ggatccaatt gtgtccattg atctctcaga gtggctgagg ataataagat ttcttcttca 240
aggtctcaag gtctgaagca tcccacagaa tgatcctact gaataactcc cataagctgc 300
tggccctata caaatccttg gccaggagca tccctgagtc cctgaagggt tatggctctg 360
tgtatcacat caatcacggg aaccccttca acatggaggt gctggtggat tcctggcctg 420
aatatcagat ggttattatc cggcctcaaa agcaggagat gactgatgac atggattcat 480
acacaaacgt atatcgtatg ttctccaaag agcctcaaaa atcagaagaa gttttgaaaa 540
attgtgagat cgtaaactgg aaacagagac tccaaatcca aggtcttcaa gaaagtttag 600
gtgaggggat aagagtggct acattttcaa agtcagtga agtagagcat tccagagcac 660
tcctcttggg tacggaagat attctgaagc tcaatgcctc cagtaaaagc aagcttggaa 720
gctgggctga gacaggccac ccagatgatg aatttgaaag tgaaactccc aactttaagt 780
atgccagct ggatgtctct tattctgggc tggtaaatga caactggaag cgagggaaga 840
atgagaggag cctgcattac atcaagcgtg gcataaga cctgccagca gcctgtatgc 900
tcggcccaga gggagtcagg gtctcatggg taacctgga cccttcttgt gaagttagaa 960
tggcctacag catggaaaaa taccgaagga caggcaacat ggcacgagtg atggtgcgat 1020
acatgaaata tetgcgtcag aagaatatc cattttacat ctctgtgttg gaagaaaatg 1080
aagactcccc cagatttgtg gggcagtttg gtttcttga ggctcctgt gagtggcacc 1140
aatggacttg ctaccacag aatctagttc cattttagac aatgaagctg cttagtaatc 1200

```

tctgccaa	gc	catctcttaa	tattaaagca	gacaccacag	aatagatttc	ttcacttaca	1260
aatgcata	tt	gggcacttat	aatacagcag	gaactcttct	cacctggagc	cttgatgtta	1320
aaagacac	ag	ccatgctctt	gaggagctta	caatcctggc	tggaggcagg	ggaggggtata	1380
ttcttttaa	at	atgcttaagt	gttatagggg	aagacggggg	taccagtaaa	catgtaacta	1440
gaaagccagg	ct	cagttctt	acctctggga	atcagaactc	tttatgcaac	ttgggttaata	1500
gaatctacta	tct	ggaagat	aaatgaagga	ttttaataaa	attttcaata	gaataaacct	1560
aatctgtatg	gata	ctttat	caaaaatgaa	tgtccctgct	atctctggat	ttatgaggca	1620
atggtacact	aa	agaatgga	atcagttcag	tgagtagaaa	ggatatccaag	gtgaagcctg	1680
agacgaatgg	cttt	cccagg	ctaccttcca	tcactgttgt	acagaaaaga	aatccagaga	1740
atcaaagtga	ctgg	ccttgg	gggtctctgc	tatggaaatg	ccattttttg	tgtctccttt	1800
ctcctactct	ttct	cacatc	ctcttcatga	ttgacgcag	gcacaaggca	agggtgtggc	1860
tgcaggtctg	gttg	aaaagt	cagcctttgg	tgtttgcaca	actgctaaca	aaaggggaca	1920
ggggaattcc	gggg	aaaatt	ttgcccctaa	ggttg			1955

<210> 62  
 <211> 1434  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1915064CB1

<400> 62							
tccgggcggtg	ccccgggctg	ggcgaggggc	cgggtgcggg	gccgctggcc	gagaggctga	60	
ggcgggcgta	tgctctccga	ggtgtccgcg	cgccgcgacg	ccaagaagct	ggtgcgctcc	120	
ccgagcgccc	tgcgcatggt	gcccgaacac	cgccgcttcg	gaagcccgtt	cggcctggag	180	
gagccgcagt	gggtcccgga	caaggagtgt	cggagatgta	tgagtggtga	cgccaagttt	240	
gactttctca	ccagaaaagca	ccactgtcgc	cgctgcggga	agtgcctctg	cgacagggtg	300	
tgacagccaga	aggtgccgct	gcggcgcatg	tgctttgttg	accccgctgc	gcagtgccgc	360	
gagtgcgccc	tggtgtccct	caaggaggcg	gagttctacg	acaagcagct	caaagtgtc	420	
ctgagcggag	ccaccttcc	cgtaacgttt	ggaaactcag	agaaacctga	aactatgact	480	
tgctgtcttt	ccaataacca	gagatacttg	tttctggatg	gagacagcca	ctatgaaatc	540	
gaaattgtac	acatttccac	cgtgcagatc	ctcacagaag	gcttccctcc	tgaggagggc	600	
aacgcacggg	ccacaggcat	gttcctgcag	tatacagtgc	cggggacgga	gggtgtgacc	660	
cagctgaagc	tgacagtgg	ggaggacgtg	actgtgggca	ggaggcaggc	ggtggcgtgg	720	
ctagtgtcca	tcacaaggc	tgccaagctc	ctctatgaat	ctcgggacca	gtaactctac	780	
gtggggctga	gcttggagta	cgtgtggtca	ccaggactga	gtcgcttgga	acagcagagc	840	
ctgtctcttg	cgtaccacag	ggattaatcc	tgttgtgtct	gggaaatgca	actcactcat	900	
gtatttggag	aaacaggagt	gttcacttat	ctagtgaat	atgttcacag	tttattaatg	960	
ctttaaacag	cttcattgtt	tagaatttgt	gtattgtcca	tacttaattg	ggggtgggag	1020	
agactgagct	acactactgc	taaactat	ttagcataat	atataccatt	tttatgagtt	1080	
cgcaggtcta	ctagaaggtt	ctggcccatc	aatattcatt	tcattttaatt	ctccacagag	1140	
accagtttgg	gcagtaggaa	ctcaggtctc	tggtctgcag	tgagcctgt	tcgcctctaa	1200	
tagccagttt	acagcacttg	ccttagcctg	tttcacagac	ttgtccactt	accttgtcac	1260	
taatttgggg	cttctgtggc	gtgagtgate	ctttgatact	tcaccaaggg	gaacgtgggg	1320	
gctttgtgtt	ttgtactttt	cactcactat	ttcactttat	taagatgact	gtacagcaat	1380	
ttgtatataa	agcttatgat	taaaaactat	tttgaacata	aaaaaaaaaa	aaaa	1434	

<210> 63  
 <211> 1360  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2185608CB1

<400> 63							
ccgataacga	acgagactct	ggcatgctaa	ctagtgcag	tcaagatggc	gggagcagct	60	
accaggctt	ccctggagtc	ggccccacgg	atcatgcggc	tggtggccga	atgcagccgc	120	
tccagggccc	gggcaggcga	gctgtggctg	cgcatggga	cagtggccac	tctgtgttc	180	
atgccagtgg	gcacgcaggc	caccatgaag	ggcatcacga	ccgaacagct	ggacgctctg	240	
ggttgcgcga	tctgcctggg	caatacctac	catctgggtc	taaggccggg	acccgagctg	300	
atccagaaa	ccaacggtct	ccacggcttc	atgaattggc	ctcataatct	gctaacggac	360	
agcggcggtt	tccagatggt	gtcgctgggt	tctctgtccg	aggtgacgga	ggagggcgct	420	
cgttccgct	ccccctacga	cggcaatgag	accctgtcga	gcccgagaa	atccgtgcag	480	

atccagaatg	cgctgggctc	ggacatcatc	atgcagctgg	acgacgtgg	tagcagtact	540
gtgactgggc	cacgtgtgga	ggaggccatg	tacaggtcaa	tccgctggct	ggaccgggtgc	600
attgcagccc	atcagcggcc	ggacaagcag	aacctcttcg	ccattatcca	gggtgggctg	660
gacgcagatc	tccgggccac	ctgccttgaa	gagatgacca	agcgagacgt	gcctggcttc	720
gccatcgggg	gcctgagcgg	gggtgagagc	aagtcgcagt	tctggcggat	ggtggcgtg	780
agcacctctc	ggctgccgaa	ggacaagccc	cgatatctga	tgggggttgg	ctatgccact	840
gatctggtag	tctgcgtggc	tcttggtatg	gacatgttcg	actgcgtctt	ccccacacgg	900
acagcgcgct	ttggctctgc	cctgggtgcc	actgggaacc	tgcagttag	gaagaagggtg	960
tttgagaagg	acttcggccc	catagaccgg	gagtgacact	gccccacgtg	ccaaaagcac	1020
agccgcgcct	tcctgcacgc	actgctgcac	agtgaaca	cggccgcgct	gcaccacctc	1080
acggtccaca	acatgcctta	ccagctgcag	ctcatgagcg	ccgtccgcac	cagcatcgtg	1140
gagaagcgct	tcccgactt	cgtgcgggac	ttcatgggcg	ccatgtacgg	ggatcccacc	1200
ctctgtccca	cctggggcac	tgacgctctg	gcctctgtgg	gaatcacact	gggctgacct	1260
ggcattggga	gagggagggg	ggaaggaagg	gagggagggg	ctggaagata	ctgaaggatt	1320
cctttttgaa	aggtttttt	tattgtaact	taaaaaaaaa			1360

&lt;210&gt; 64

&lt;211&gt; 2277

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2228862CB1

&lt;400&gt; 64

aaacgccatg	cggagggggc	agcgcaggga	cgccggacgt	ccgcgcccg	agtccccgg	60
gcccgcgggc	agggcctcgc	tggaggagcc	gcctgacggg	ccgtctgccg	gccaagccac	120
cgggcccggc	gagggccgcc	gcagcaccga	gtccgaggtc	tacgacgacg	gcaccaacac	180
cttcttcttg	cgagcccaca	ccttaaccgt	gctcttcac	ctcacctgta	cgcttggcta	240
tgtgacgctg	ctggaggaaa	cacctcagga	cacggcctac	aacaccaaga	gaggtattgt	300
ggccagtatt	ttggttttct	tatgttttgg	agtcacacaa	gctaaagacg	ggccattttc	360
cagacctcat	ccagcttact	ggaggtttgg	gctctgcgtg	agtgtgggtc	acgagctgtt	420
tctcatcttt	atactcttcc	agactgtcca	ggacggccgg	cagtttctaa	agtatgttga	480
ccccaaagctg	ggagtcccac	tgccagagag	agactacggg	ggaaactgcc	tcacttacga	540
cccagacaat	gagactgacc	cctttcacaa	catctgggac	aagttggatg	gctttgttcc	600
cgcgcacttt	cttggctggg	acctgaagac	ctcatgatc	cgagactggg	ggatgtgcat	660
gatcatcagc	gctgatgttc	agttcctgga	gtacagcctg	gagcaccagc	tgcccaactt	720
cagcgagtgc	tgggtgggac	actggatcat	ggacgtgctc	gtctgcaacg	ggctgggcat	780
ctactgcggc	atgaagacct	ttgagtggct	gtccctgaag	acgtacaagt	ggcaggccct	840
ctggaacatt	ccgacctaca	agggcaagat	gaagaggatc	gccttcag	tcacgcgta	900
cagctgggtt	gccttcgagt	ggaagccggc	ctccagcctg	cgtcgtggc	tgcccggtg	960
cggcatcatc	ctggtgttcc	tgttggcaga	actgaacacg	ttctacctga	agtttgtgct	1020
gtggatgccc	cggagcact	acctggtcct	cctgcggctc	gtcttcttcg	tgaacgtggg	1080
tggcgtggcc	atgctgtaga	tctacgactt	catggatgac	ccgaagcccc	acaagaagct	1140
gggcccgacg	gcctggctgg	tggcgcccat	cacggccacg	gagctgctca	tcgtggtgaa	1200
gtacgacccc	cacacgctca	ccctgtccct	gcccttctac	atctcccagt	gctggaccct	1260
cggctccgct	ctggcgctca	cctggaccgt	ctggcgcttc	ttcctgcggg	acatcacatt	1320
gaggtacaag	gagaccgggt	ggcagaagtg	gcagaacaag	gatgaccagg	gcagcacctg	1380
cggcaacggg	gaccagcacc	cactggggct	ggacgaagac	ctgctggggc	ctgggggtggc	1440
cgagggcgag	ggagcaacaa	ctccaaactg	acctgggccc	tggctgcctc	gtgagcctcc	1500
cagagcccag	gcctccgtgg	cctcctcctg	tgtgagtcct	accaggagcc	acgtgccccg	1560
ccttgccctc	aagggttttt	gcttttctcc	tgtgcacctg	gcgaggctga	aggcgagggg	1620
tggaggaggc	cccagcacag	cctcatctcc	atgtgtacac	gtgtgtacgt	gtgtatcgct	1680
gtgtgtacgc	gtgtgtacgc	gcgtgtgtac	acatgcgtgg	ccgcctgtgg	tgtgcacgtg	1740
tgctctgggc	tccgaggctt	ctccagagct	gggagctggc	tggcgtggca	agggcattgct	1800
ctggggcagt	gtgtccctca	ggaaccaggg	tcctccctcc	cctttctgcc	tggtcagccc	1860
cgtggcctct	ggcccaccaa	gtccctgtc	acctagccat	ggtgtgggtc	aggcagggac	1920
atctcggtag	ctttctgca	ctcctggggc	cctgggtgcg	ctgaggcctg	gaggcgctca	1980
cactggctcc	acatccactt	cccccgagc	tcgtgtgggc	gtcgtccac	aaacactccg	2040
tggctgagag	gcagcggatc	caggcagcga	tgctgagcca	cctcctccga	gccttctctt	2100
caacagagac	accccgagg	acacgtggat	catgggggtca	gagatcactg	agctgccccct	2160
caagggggccc	tggaacccgg	gtgctgggtg	catgtgcct	ccgtggctcc	aaggtagggg	2220
tcactcttcac	gagcaaaagag	aaccaataaa	gtgacaacga	acgtcaaaaa	aaaaaaa	2277

&lt;210&gt; 65

&lt;211&gt; 1592

<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2235577CB1

<400> 65  
ccaagatggc ggcccccgagg agctgtgctc tatggagcta ttggggccgt ggggtggctgc 60  
gggcatgctg gggctgccag ctctctgggc ttcttagctc ttggggccgg gacctactaa 120  
gtgctcgggt cttgtcccaa gagaagcggg cagcggaac gcactttggg ttgagactg 180  
tgctcgaaga ggagaagggg ggcaagaac tcctttaatg tgaaagtgtt tgcctggcatc 240  
acttctcttg ggacatcttc aacctttttg tcacagctcc atttcttggt catgttgtct 300  
atcaggtgtt tgaaagtgtg gctaagaagt atgatgtgat gaatgatatg atgagtcttg 360  
gtatccatcg tgtttggaag gatttgctgc tctggaagat gcacccgctt cctgggaccc 420  
agctgcttga tgttgctgga ggcacaggtg acattgcatt cgggttcctt aattatgttc 480  
agtcccagca tcagagaaaa cagaagaggc agttaagggc ccaacaaaat ttatcctggg 540  
aagaaattgc caaagagtac cagaatgaag aagattcctt gggcgggtct cgtgctgctgg 600  
tgtgtgacat caacaaggag atgctaaagg ttggaagca gaaagccttg gctcaaggat 660  
acagagctgg acttgcatgg gtattaggag atgctgaaga actgcccttt gatgatgaca 720  
agtttgatat ttacaccatt gcctttggga tccggaatgt cacacacatt gatcaggcac 780  
tccaggaagc tcctcgggtg ctgaaaccag gaggacggtt tctctgtctg gaatttagcc 840  
aagtgaacaa tccccctcata tccaggcttt atgatctata tagcttccag gtcacccctg 900  
tcttgaggga ggtcatcgct ggagactgga agtctctatca gtaccttgta gagagtatcc 960  
gaagggttcc gtctcaggaa gagttcaagg acatgataga agatgcaggc ttccacaagg 1020  
tgactttacga aagtctaaca tcaggcattg tggccattca ttctggcttc aaactttaat 1080  
tcctttccta tcattggagca tgaaccagtc atatcctggt gaaagccttg aactgaagga 1140  
taatctggca aatgagacag cagcagagca tctctcttta aggatacgtg ccttggtactc 1200  
atgtttgaat cgaacagctc caaagtggaa gaacaaattc ttgtcacttt ttacagctt 1260  
tctttggagc tgccttcagtc catctccag aggcatttgg tctgtatctt tgctcaactg 1320  
ctaatttctc ttggctgtag ggtgtgtggt taaggtaaaa ccacccctaa agctcagttt 1380  
tgaagttagt gtatttatag cttctctgct ggtgctgcct tctagaggga tgatagatca 1440  
tttgaaccca atgacaattt ttaaccagaa aacctttcag gttgtgctgaa 1500  
cctaggagca agtctaggcc caagtcagag tattaatgat catgagaatt gtgtgctgaa 1560  
ccagtaaacg agttttacctt ttaaaaaaaa aa 1592

<210> 66  
<211> 1390  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2271680CB1

<400> 66  
gatcggttgt cgcgatgtgg agtggccgta agctgggctc ctccgggggt tggtttttaa 60  
gagtgctggg gcctggaggc tgtaatacaa aagctgcgcg tcctttaatt tcctcggcgg 120  
tttatgtgaa gaaccagctc agtgggactc tacagattaa accagggggt ttcaatgaat 180  
acagaacat atggttcaaa tcctacagga cgatcttttc ctggttgaaac agaataaaga 240  
gtttcagggt gagtttcaat tctgttgccc aggcctggagt acaatgggtg gatcttggct 300  
cactgcaacc tccacctccc ggttccaage gattctcctg cctgagcctc ctgagtcatt 360  
gggattacag gtacctcttg gcgagactgt acagtacttc ccaaaccact gtcgacagcg 420  
gtgaggtaaa aacctctctg gccctggctc acaaatgggt ggatgaacaa ggagtatatg 480  
cacctcttca ttccatgaat gacctgaggg tgccatttat tagggacaat cttctgaaaa 540  
caattcctaa tcaccagcca ggaaaacctt tgttggggat gaagattctt gacgttggct 600  
gtggtggtgg gctgttaact gaacctctag ggcggcttgg ggcttcagtt attggaatcg 660  
accctgtgga tgagaacatt aaaacagcac aatgccataa atcatttgat ccagtcctgg 720  
ataagagaat agagtacaga gtgtgttccc tggaaagat tgtggaagag actgcagaaa 780  
catttgatgc tgttgtagct tctgaagtgt tagaacatgt gattgatcta gaaacatttt 840  
tacagtgcct ctgtcaagtg ttaaaacccg gtggttcttt attcattact acaatcaaca 900  
aaacacaact ttccatagcc ttgggaattg ttttttcaga gcaaatgca ggtattgtac 960  
caaaaggtac tcatacatgg gagaagtgtg tttcaactga aacactagag agcattctgg 1020  
aatcaaatgg tctgtcagtt caaacagtg taggaatgct ctataacccc ttctcaggtt 1080  
actggcattg gagtgaaaat accagcctta actatgcagc tcatgctgtg aaatccaggg 1140  
tccaggaaca cccagcctct gctgagtttg ttttaagggt agaaacagaa gagctccaa 1200  
ctaattgcctg caccaatcca gctgtgcatg aaaagctgaa gaaatgaatt gttcttgaga 1260

actatagtaa tatggcttgg atatctgatg ttttcaaata caagaaatgt acaattttatc 1320  
ctttgagaga gaatcatgaa gaaaagaagg tcaataaaaaa gggctaaaaac cttggaaaaa 1380  
aaaaaaaaa 1390

<210> 67  
<211> 899  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2325603CB1

<400> 67  
ttcattggcg ggtccgccgc ggtcggaggat cctggaggtc ttaaccatga acttctcttg 60  
aggagggagg caggaagcag caggggccag gagtagaagg gctcccagac cccgagaaca 120  
ggaccgagac gtgcagctgt ccaaggctct gtcctatgcc ctgcccagat gggccttgaa 180  
gctggggctt cccatgggag ctgatggctt cgtgcccctg ggcaccctcc tgcagttgcc 240  
ccagttccgc ggcttctctg ctgaagatgt gcagcgcgtg gtggacacca ataggaagca 300  
gcggttcgcc ctgcagctgg gggatcccag cactggcctt ctcacccggg ccaaccaggg 360  
ccattcccct caggtaacct agttggagct gatgccctg gagacaccgc aggcctgcc 420  
cccgatgcta gtccatggta cattctggaa gcactggcca tccatcctac tcaaaggcct 480  
gtcctgccag ggaaggacgc acattcacct ggccccagga ctgctgggag accccgggtat 540  
catcagtgga atgcgggtccc attgtgaaat agctgtgttc atcgatggac ccttggtctt 600  
ggcagatgga atacccttct tccgctctgc caatggggtg attctgactc cagggaatac 660  
tgatggcttc ctcttcccca agtacttcaa ggaggccctg cagctacgcc ctaccgaaa 720  
gccctttcc ttggctgggtg atgaagagac agagtgtcag agtagccca agcacagctc 780  
cagagaaagg aggaggatcc aacaataaaa tattaattta taaaaagaa attttaaaaa 840  
gtaacaagaa agaactcggt tgaaccatg tttcatcatc ctgtaaaaa aaaaaaaa 899

<210> 68  
<211> 1041  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2356055CB1

<400> 68  
gacgtaggag ctgatagcaa ggagggggca catttcaggg acaccatgaa ggggtggttc 60  
actgggggtg atgagtacca gaagcacttc ctgccaggg actacttggc tacttactac 120  
agcttcaatg gcagcccctc acccgaggcc gagatgctga agtttaactt ggaatgtctc 180  
cacaagacct tcggccctgg aggcctccaa ggggacacgc tgattgacat tggtcaggt 240  
cctaccatct accaagtctt tgctgcctgt gattccttcc aagacatcac tctctccgac 300  
ttaccggacc ccaaccggga ggagctggaa aagtggctga agaaggagcc gggggcctat 360  
gactggaccc cagcgggtgaa attcgctgtg gagctggaag gaaacagcgg ccgatgggag 420  
gagaaggagg agaagctgcg ggcagcgggtg aagcgggtgc tcaagtgcga tgtccacctg 480  
ggcaaccgcg tggcccgccg tgtgttgctt ctgcgcgact gtgtgtctac cctgtggcc 540  
atggagtgtg cctgctgtag ccttgatgcc taccgctgt cctgtgcaa ccttgctca 600  
ctgctcaagc cgggtggcca cctggtgacc actgtcacgc ttcggctccc gtcctacgtg 660  
gtggggaagc gtgaattttc ctgctgggcc ctggagaaag aggaggtggc agccaggcaa 720  
tgtccaggag aggagattgc caaggaaagg aggcataaaa tgccccctcc ttgtgatgtc 780  
aggacctccc ttagecagcg atctggccaa gacacaggga aaagacacag gatccagacc 840  
cggggctctg ctcttgggac ggctcagtcg agagagtcag ctggtgcctt ggaaggagag 900  
agtcggcaag ggtgtgaggg aatctttggg tgctgtggaa gctgttctac cttatgaaat 960  
ggggctggga tggactgagt gactatgctg tgctctgtca tttgtccgta agtactcgct 1020  
gcatactctg atgcgctgat g 1041

<210> 69  
<211> 1106  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2448909CB1



&lt;400&gt; 69

```

acggccagtg caagctaaaa ttaaccctca ctaaagggaa taagcttgcg gccgccgcgc 60
ctagaactgt atttcagaaa aaagaaacta cagtttttagc atgcagaaag gaaaagggag 120
aacaagccgg atcagaagac gaaaactctg cggaagtctt gaatcaagag gagtgaatga 180
gagccacaag tctgaattta tagagctgag gaagtggctg aaagctagga agtttcaaga 240
ttcaaaactta gcgcctgctt gttttccagg tacaggaaga gggctgatga gtcaaacatc 300
cctgcaggag ggacagatga ttatttcggt gcctgagagt tgccctgctca ccacggacac 360
agtattcgga agctacttag gggcatacat tactaagtgg aagcctcctc catctcctct 420
gctggcgctg tgcacctttt tagtttcaga aaagcatgct gggcaccgat ctctttggaa 480
gccttacctg gagattttac ccaaggcgta tacctgccct gtttgtttg agccggaagt 540
ggtgaacctt ctcccaaat ctttaaaagc aaaggctgaa gagcagagag cccacgtgca 600
ggagtctttt gcttcctcca gagacttttt ctcttctctg cagcctctgt ttgaggaggc 660
tggtgacagc atcttcagct acagtgcctt gctgtgggct tgggtgcaccg tcaacaccag 720
agccgtgtac ctgaggccca ggcagcgagg atgcctttct gcagagccgg acacctgtgc 780
actcgctccg tacctggacc tgctgaatca tagccacat gtccaggtaa aagcagcggt 840
taatgaagaa actcattctt acgaaattag aacgacttca cgttgagaa agcatgaaga 900
ggtattcatc tgttacggcc ctcacgataa tcaacggctg ttcttggaa acggatttgt 960
ttctgtccat aatcctcatg cttgtgttta tgtctcaaga ggttggaate aactttgttc 1020
ttaacattaa cactatataa tttttttccc catttgagaa tgtgtatttt cagttttaat 1080
aaaaatatca aaaccttaaa aaaaaa 1106

```

&lt;210&gt; 70

&lt;211&gt; 2405

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2631212CB1

&lt;400&gt; 70

```

acgggggtct ggctggcgcc agcgggcgga gggagccgag agaccgagtg gcacgtgtgg 60
agaagcggcg gcacaagcgc ggcgggcgga gacactcccg cccccaccag actcaagccc 120
tcactcgact ctgcggcctc tegtgtctcg cacagctccc tgcccaggct aggaggccgg 180
cttcgggggt tgagtggccc gagctaaggg tgccgagacc taaggcgccg gactacgacg 240
cggttgatat cgggtgtaac gacggcctca gcaggcgggg aagatgaaag gtacccggat 300
cgagctggga cagtgcgac cacacaatat taaacagttg aaaagattga atcaggtcat 360
ctttccagtc agctacaatg acaagttcta caaggatgtg ctggaggttg gcgagctagc 420
aaaacttgcc tatttcaatg atattgctgt aggtgcagta tgctgtaggg tggatcattc 480
acagaatcag aagagacttt acatcatgac actaggatgt ctggcacctt accgaaggct 540
aggaaatagga actaaaatgt taaatcatgt cttaaacatc tgtgaaaaag atggtacttt 600
tgacaacatt tatctgcatg tccagatcag caatgagtcg gcaattgact tctacaggaa 660
gtttggcttt gagattattg agacaaagaa gaactactat aagaggatag agcccgca 720
tgctcatgtg ctgcagaaaa acctcaaagt tccttctggt cagaatgcag atgtgcaaaa 780
gacagacaaa tggcctttta acaaatgaac tttctgcac ttgctgtcg ccaataaaaa 840
gagaggccca ttgattcctc cccaccccca acacttttct tttaaagctt ttctccctcc 900
ttgttcttgt ttttctttct tcctttcctt ttctctgaga gttttaatac tttcaaggac 960
tttaaaaaaa taatcatgtt tgaattgttt tctcttattt ttgtgaggtg gtttgaagga 1020
aggacaaggt agatctgttt agttttgcag ttgaagttag atggtcctaa acatttaatt 1080
gtcaataaat ttcaaattta atgtcctgct ttcacattga agggcagagc ctacaaaaca 1140
ttgtatattt caaaagacaa aaagaagcag cagcagatc ttgttctcta attcatagac 1200
aagttgagtg tgtttgtggt actttggggt tttaaacact ttgggatact aatccctaga 1260
cattgccttc actccacctt tagtcttctc gagcactctc tcgggagttg gaacatttgt 1320
atccttgtaa gaaataactaa gcttatgttg atttttaagt aattatatct tctcttcttg 1380
ctggtgggtg gggcagtttg gtttagtggt atactttggt ctaagtattt gaggtaaaact 1440
gcttttttgc taatgagtggt gctggttggt agcaggtttg tttttcctgc tgttgattgt 1500
tactagtggc attaactttt agaatttggt ctggtgagat taattttttt taatatccca 1560
gctagagata tggcctttta ctgacctaaa gaggtgtggt gtgatttaat tttttccogt 1620
tcctttttct tcagtaaac caacaatagt ctaaccttaa aaattgagtt gatgtcctta 1680
taggtcacta cccctaaata aacctgaagc aggtgttttc tcttgacat actaaaaaat 1740
acctaaaagg aagcttagat gggctgtgac acaaaaaatt caattactgt catctaattgc 1800
cagctgttaa aagtgtggcc actgagcatt tgattttata ggaaaaaata gtatttttga 1860
gaataacata gctgtgctat tgcacatctg ttggaggaca tcccagattt gcttatactc 1920
agtgcctgtg atattgagtt taaggatttg aggcaggggt aattattaaa catattgctt 1980
ctattcttgg aaaaatagaa gtgtaaaatg ttaataatac aaatgtcact gtgacctcct 2040
ccactgagag gactggttta tgccagatca ttttccggca cacacggagt ggctttgaca 2100
gattgataac tttgtaagat gggagacatc tgaatatatt atgttttctt tttgtagtcc 2160

```

```
catctccact atttagaaat gttctcagac tttaaaataa tgcacagggc ttgagctttc 2220
tgtcatttga cttttaaagg aagtttcatt catattttat ctcttatgta aaattgcggt 2280
ataaagtctc atttccaaat atgttaaattg acaaaattat tttataaaat gtttatgcac 2340
actttataac ctttaagtttt tatttgagaa tgtgaaagta caaagtgcag tagacttcaa 2400
caatc 2405
```

<210> 71  
<211> 1239  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2678733CB1

```
<400> 71
ggcgatgagg gtgttggtgc ggcgctgttg gggctcctccg ctggctcatg ggcgccaggcg 60
tgaggaggcg agtccccagt ggcgagcact ggcccagact ggctgggagg actgccggga 120
ctccagagtc cgcgagaagc ctccctggcg ggtgctcttc ttcggcacgg accagttcgc 180
ccgcgaggcg ctgcggggcg tgcacgccgc cagggaaaac aaagaagaag agttaatcga 240
caaacaggag gtggtcacaa tgccttcccc atcacaaaaa ggactgccag tgaagcaata 300
tgctgtgcag tctcagcttc ccgtatatga gtggccggat gtgggatctg gagaatatga 360
tggtggagta gtggttcgt ttggccgact ttggaatgag gctcttattc ttaaatttcc 420
ctatggcata ttgaatgttc atcccagttg cctcccagaga tggcgtggcc cagccccctgt 480
aatccataca gtggttcacg gagacacagt tactggagta acaattatgc aaattagacc 540
taaaagggtt gatgtaggcc caattctcaa acaagaaact gttcctgtgc cacccaagag 600
cactgcaaag gaattggaag cagtgtgtgc aagactgggt gcccaacatgc tcatttcagt 660
tttgaaaaat ttgcctgaaa gtctgagcaa tggaaggcag cagccaatgg agggggcgag 720
ttacgccctt aagatttctg ctggtaccag ttgtataaaa tgggagggaac aaacttcaga 780
acaaatattc agactttacc gtgccattgg aaatataatt ccgttgcaga cgctctggat 840
ggcgaatacc attaaacttc tggatttggg agaagttaac agttcagtc ttgctgatcc 900
aaaattaacg ggacaggctc ttattccagg atcagtaata taccacaaac agtcacaaat 960
actattgggt tattgcaagg atggttggat ttgtgttcga tcagtgtatgc tcaagaaatc 1020
actaacagct actgacttct acaatggata tttgcacccc tgggtaccaga aaaattccca 1080
agctcaacca agccaatgca gatttcagac tctcagactt ccaacaaaga agaagcagaa 1140
aaaaactggt gctatgcaac aatgcattga gtagttagga agaagatgga taaaaacct 1200
ttacatattt gtaatttatt aaaaacctta ttacaagg 1239
```

<210> 72  
<211> 2295  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2768571CB1

```
<400> 72
aaactcagca cttgccggag tggctcattg ttaagacaaa. ggggtgtgcac-ttcttgcca-60
ggaaacctga gcggtgagac tcccagctgc ctacatcaag gcccaggac atgcagaacc 120
ttctcttaga acccgaccca ccaccatgag gtccctgctg tggagatgca ggcacctgag 180
ccaaggcgtc cagtggctct tgcctctggc tgcctctggtc ttctttctct tcgcttgcc 240
ctcttttatt aaggagcctc aaacaaagcc ttccaggcat caacgcacag agaacattaa 300
agaaaggctt ctacagtccc tggcaagcc taagtcccag gcaccacaa gggcaaggag 360
gacaaccatc tatgcagagc cagtgccaga gaacaatgcc ctcaacacac aaaccagcc 420
caaggcccac accaccggag acagaggaaa ggaggccaac caggcaccgc cggaggagca 480
ggacaagggtg cccacacag cacagagggc agcatggaag agcccagaaa aagagaaaac 540
catggtgaac acactgtcac ccagagggca agatgcagg atggcctctg gcaggacaga 600
ggcacaatca tggaagagcc aggacacaaa gacgacccaa ggaaatgggg gccagaccag 660
gaagctgacg gctccagga cgggtgtcaga gaagcaccag ggcaaagcgg caaccacagc 720
caagacgctc attcccaaaa gtcagcacag aatgctgggt cccacaggag cagtgtcaac 780
aaggacgaga cagaaaggag tgaccacagc agtcatccca cctaaggaga agaaacctca 840
ggccaccca cccctgccc ctttcagag cccacgacg cagagaaaacc aaagactgaa 900
ggccgccaac tcaaatctg agcctcgggt ggattttgag gaaaaataca gcttcgaaat 960
aggaggcctt cagacgactt gccctgactc tgtgaagatc aaagcctcca agtcgctgtg 1020
gctccagaaa ctctttctgc ccaacctcac tctcttctct gactccagac acttcaacca 1080
gagtgagtgg gaccgcctgg aacactttgc accaccctt ggcttcattg agtcaacta 1140
```

ctccttggtg	cagaaggtcg	tgacacgctt	ccctccagtg	ccccagcagc	agctgctcct	1200
ggccagcctc	cccgctggga	gcctccggtg	catcacctgt	gccgtggtgg	gcaacggggg	1260
catctgaac	aactccaca	tgggccaggga	gatagacagt	cacgactacg	tggtccgatt	1320
gagcggagct	ctcattaaag	gctacgaaca	ggatgtgggg	actcggacat	ccttctacgg	1380
ctttaccgcc	ttctccctga	cccagtcact	ccttatattg	ggcaatcggg	gtttcaagaa	1440
cgtgcctcct	gggaaggacg	tccgctactt	gcacttcctg	gaaggcacc	gggactatga	1500
gtggctggaa	gcactgctta	tgaatcagac	ggtgatgtca	aaaaaccttt	tctggttcag	1560
gcacagaccc	caggaagctt	ttcgggaagc	cctgcacatg	gacagggtacc	tggtgctgca	1620
cccagacttt	ctccgataca	tgaagaacag	gtttctgagg	tctaagaccc	tggatgggtgc	1680
ccactggagg	atataccgcc	ccaccactgg	ggccctcctg	ctgtcactg	cccttcagct	1740
ctgtgaccag	gtgagtgttt	atggcttcac	cactgagggc	catgagcgct	tttctgatca	1800
ctactatgat	acatcatgga	agcggctgat	cttttacata	aaccatgact	tcaagctgga	1860
gagagaagtc	tggaagcggc	tacacgatga	agggataatc	cggtgtgacc	agcgtcctgg	1920
tcccgggaact	gccaaagcca	agaactgacc	ggggccaggg	ctgccatggt	ctccttgcc	1980
gtccaaggc	acaggataca	gtgggaatct	tgagactctt	tggccatttc	ccatggctca	2040
gactaagctc	caagcccttc	aagagttcca	agggaaact	tgaaccatgg	acaagactct	2100
ctcaagatgg	caaattggcta	attgaggttc	tgaagttctt	cagtacattg	ctgtaggtcc	2160
tgaggccagg	gatttttaat	taaattgggt	gatgggtggc	caataccaca	attcctgctg	2220
aaaaacactc	ttccagtcga	aaagcttctt	gatacagaaa	aaagagcctg	gatttacaga	2280
aacatataga	tctgg					2295

&lt;210&gt; 73

&lt;211&gt; 2182

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3189062CB1

&lt;400&gt; 73

ccggcatggc	cggtgagct	gcaggcttcc	ttatttaaga	ccgggaattt	agtcagcctg	60
gtgagccgac	tctgaggaga	tgagtatcgc	ctgaggtgat	gagagagaat	gtggtgttta	120
gcaacatgga	gagagaaagt	gggaagcccg	tggtgttgt	cgcagtgtgt	actgagcctt	180
ggtttaccca	gcgatacaga	gaatatctcc	agaggcagaa	actctttgat	acacagcacc	240
gtgtggaaaa	gatgccggat	ggctcgggtg	cgctaccggt	gctgggagag	acgcttcag	300
agcagcacct	gcaggagctg	aggaatcgtg	ttggcccagg	cagtccctgt	atgtcacgc	360
agctcccga	tctgttctt	tgaagagggt	cccagggttg	ttcacctgcc	caaaaattgt	420
gtcttgaggt	gagtcgctgg	gtggtgggtc	ggggagtcac	gtggtcagcc	gagttggagg	480
tctgatttgc	ccgatcatgg	caacggcatg	gtaatctctt	gttctgagt	gaagactgtt	540
ccgaagccaa	ccagtggaaa	aatctgggac	cggaactctg	ggagaccgtt	gccttggcac	600
ttggcgteca	gcgtttggca	aaacgagggc	gggtatcacc	ggatgggtact	cgaactccag	660
cagtgcact	gctgctgggt	gaccatggct	gggtagagca	tgtggataat	ggtatccgtt	720
ataagtttga	cgtgaccag	tgtatgttct	cctttggaaa	catcactgag	aagcttcag	780
tggtcctg	gtcctgtgct	ggagaatgct	tggtggatct	ctatgcagg	attggttatt	840
ttacattgcc	tttctagtt	catgctgggt	ctgccttcgt	ccatgcttgt	gagtggaa	900
cccattgctg	agttgctctg	agaaataacc	ttgagatcaa	tggagtagca	gatcgggtcc	960
aaatacactt	tggagataac	agaaaactga	agctctcaaa	tattgcagat	agggtgatcc	1020
tgggctgat	tccagctct	gaagaaggct	ggcccattgc	ctgccaaagt	ttaaggcagg	1080
atgctggagg	cattttgcat	atccaccaaa	atgtgggaatc	tttcccagg	aagaatcttc	1140
aggctcttgg	agtcagcaaa	gtagagaaag	agcattggct	gtatcctcag	caaattacca	1200
ccaaccaatg	gaaaaatgga	gctaccagg	attctagggg	aaaaatgctg	tcaccagcca	1260
ccaagccaga	gtggcaaagg	tgggcagaat	ctgcagaaac	tcgaatcgcc	actcttcttc	1320
agcaggtgca	tggaagacac	aaattctgca	catccaacca	gtgaaatcct	gtgaaatcct	1380
atgctcccca	tgtggatcac	atagtcctgg	atctggaatg	ctgcccctgt	ccttcagttg	1440
gctagaggag	gtagatcctg	ggacacatgg	gatccacgtg	cgagtggccc	ttaaattgat	1500
cagttcagtc	caggttgtea	tcccttttgt	cccctgggtg	tcagtttttt	tcatatttta	1560
tagccctgaa	agcaggctct	agatcaattc	aaattatttc	atttgtcttt	cattgataac	1620
agaaaaatgaa	atacctgttt	gggagaagca	gcatggccca	ttgaaatgag	gctcatctgt	1680
gcaattatga	attccaaatt	ctgacctcag	ttctgggaatt	gaagtttcag	tatgttttgg	1740
cctcgggttt	cgttatttgc	aaaatgagag	tttctttgaa	ctgtctcag	tgactattaa	1800
gcaactatac	acaggacac	ggttatttta	gagtgaagaa	cacagtgctt	tttccaaatt	1860
gctctggcta	ccatatagaa	aattgactga	aggaggggcca	agatggaaac	agagagacca	1920
gtgaggaggc	ttctgtggtt	gtccaggtct	gaggtgatgg	taacttgga	tcggatgggt	1980
gtaattggag	gtagattgat	atgataaata	aaattgacag	accaagcaat	ggaatcagaa	2040
ttaagtccct	aacatgaagc	tgctttgtta	ttatgactga	ctaaattaga	gagagaagg	2100
aacaaaaatt	attttagtg	atttccattc	ccttactgtg	gcatttctaa	atgattgtga	2160

ggggttggtct tataaatttg gt

2182

<210> 74  
<211> 1288  
<212> DNA  
<213> Homo sapiens  
  
<220>  
<221> misc\_feature  
<223> Incyte ID No: 3243884CB1

<400> 74  
cggtttctttc ggagcggcgg tgaaggtcct ggggtgaggtg ggggtggatg gtgcttgccg 60  
cgtatcatgg ctgcctccgg aaaactcagc acttgccgtc tccctccgtt gccacgatt 120  
cgagaaatca ttaagtgtt aagactgcaa gcagcgaagc agctatcaca gaatttcctc 180  
ctggacttga ggctgacaga taagattgta aggaaagctg gcaatctgac aaatgcttat 240  
gtttacgaag tggccttttg gccaggggga atcacaagat ctattcttaa tgccgacgtc 300  
gctgaacttc tgggtggttga aaaggacact cgatttatc ctggattaca gatgctttct 360  
gatgcagcac ctgggaaact gagaattggt catggagatg tcttgacatt taaggtagaa 420  
aaggcttttt cagaaagtct taaaagaccc tgggaagatg atcctccaaa tgtacatat 480  
attggaatc tgccttttag tgtttcaact ccaactgatta tcaagtggct tgaaaaatatt 540  
tcctgtagag atggaccttt tgtttatggc agaactcaga tgactttgac ttttcaaaag 600  
gaagtggcag agagacttgc agccaataca ggaagcaaac agcgtagtcg cctctctgtt 660  
atggctcagt acctctgcaa tgttcgacac atctttacaa ttccaggaca agcttttgct 720  
cccaaaccaag aggtggcagt gggcgtggtg cacttcactc ccttgataca gcccaagata 780  
gagcagccat tcaagctggt ggaaaaagtgt gttcagaatg tatttcagtt ccgaaggaaa 840  
tactgccatc gagggctcag aatgtttatc cctgaagcgc agcgttgga aagcacgggc 900  
aggctgttag agttggcaga catagacctt actcttcggc cccgccagct ctccatctca 960  
cactttaaga gcctctgtga tgtatacaga aaaatgtgtg atgaagacct acaactcttt 1020  
gcatataatt tcagagaaga actcaagcga agaaaaagca aaaatgaaga aaaagaagag 1080  
gatgacgcag agaattacag actctagctg ctgcctgggg gcgagcagcc taccagatgt 1140  
cgatttgcac tacgtggagc ttcttatata ggtactcttt tgtctttaca gaatgacgat 1200  
acaaatgcca atgaccagat gtgacttatt ttctttttac tatacagctt ggagagaaa 1260  
ataaatatca tcaataaaga aaaaaaaa 1288

<210> 75  
<211> 1130  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 3400578CB1

<400> 75  
ggtgaaccgg cagccggact tctgctgcac tggggctccg aatgaccag aatctggggt 60  
gggcttaggc cccgggagtg agatggcctc aatcttgcca acgcctcagg ctctccagct 120  
cactctagcc ctgatcaagc ctgacgcagt cgcccatcca ctgattctgg aggctgttca 180  
tcagcagatt ctaagcaaca agttcctgat tgtacgaatg agagaactac tgtggagaaa 240  
ggaagattgc cagaggtttt accgagagca tgaagccggg ccaatccgag cctacatcct 300  
tgcccaacaag gatgccatcc agctctggag gacgctcatg ggaccaccca gagtgttccg 360  
agcagcccat gtggcccagc attctatccg tgggagtttc ggccctcactg acaccggcaa 420  
caccacccat ggttcggact ctgtggtttc agccagcaga gagattgcag cettcttccc 480  
tgacttcagt gaacagcgtt ggtatgagga ggaagagccc cagttgcgct gtggccctgt 540  
gtgctatagc ccagagggag gtgtccacta tgtagctgga acaggaggcc taggaccagc 600  
ctgatgcagg tctatgaaga ccagtggttag tgcccagact tctcctagac atctagtcta 660  
aaacattctc ctaggaccag ggaagcctgg cttacagtgc catttctgct gggcaccacc 720  
acctgacctg gggcctagct caccacagca catcctccag gatctagcct tctatctacc 780  
tcttctctgg aatgtttatg gtggttcaga agaattgatga ctctctttg ctgagaactg 840  
ttcatccttc ttcaagaaga agcttgccag gccgggcacg gtgctcacgc ctataatccc 900  
agcacttttg gaggccgagg cgggcccagc acaaggctcag gaattcgaga ccagcctgac 960  
caacatggtg aaaccccatc tctactaaaa atacaacaat tagccaggca tgggtgtgca 1020  
tgcttgaat cccagctact cagaggtga ggcaggagaa ttgcttgatc ctgggaggca 1080  
gaggggtgcag tgagccgaga tcgtgccatt gcaactccagc ctgagggagg 1130

<210> 76  
<211> 1815

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3422577CB1

&lt;400&gt; 76

gacggcactg	gggtggggccg	agctccaggg	ctggctgctg	ggctgctaag	ggaactgtga	60
gccgctcaga	gccgcgcgcc	tcccggggcgg	ggcggggccg	gccgtgggag	tccgcgcgtg	120
cccgcgcgca	gctgcctgct	ccggcgggctt	cgctgctagc	tcgcggcgac	gtcggggccga	180
ttttccagg	atgacagagc	tgaggcagag	gggtggcccat	gagccgggtg	cgccaccgca	240
ggacaaggag	tcagagtcag	aagcaaaagg	agatggagag	actgcatcgg	acagtgaag	300
ccgggcagaa	tccgcacccc	tgccagtctc	tgcagatgat	accccgagg	tcctcaatag	360
ggccctttcc	aacttgtctt	caagatggaa	gaactgggtg	gtgagaggca	tcctgacttt	420
ggccatgatt	gcatttttct	tcatcatcat	ttacctggga	ccaatgggtt	tgatgataat	480
cgtgatgtgc	gtttgttcca	agtgtttcca	tgagataatc	actattggct	acaacgtcta	540
ccactatat	gatctgcctt	ggttcaggac	gctcagctgg	tactttctcc	tgtgtgtaaa	600
ctattttctc	tatggtgaga	cagtgcagga	ttacttcttc	accctgggtc	agagagaaga	660
gcctttgagg	attctcagta	aataccaccg	gttcatttcc	tttactctct	atctaataag	720
attctgcatg	tttgtactga	gtctgggtcaa	gaagcattat	cgactgcagt	tctacatgtt	780
tggtgggacc	catgtgacat	tgctgattgt	tgtaacacag	tcacatcttg	ttatccacaa	840
cctatttgaa	ggaatgatct	ggttcattgt	ccccatatct	tgtgtgatct	gtaatgacat	900
catggcctat	atgtttggct	ttttctttgg	tcggacccca	ctcatcaage	tgtccccgaa	960
gaagacctgg	gaaggcttca	ttggggggctt	ctttgctact	gtgggtgttg	gccttctgct	1020
gtccatgtgt	atgtccgggt	acagatgctt	tgtctgccct	gtggagtaca	acaatgacac	1080
caacagcttc	actgtggact	gtgagccctc	ggacctgttt	cgcctgcagg	agtacaacat	1140
tcctggggtg	atccagtcag	tcattgggtg	gaaaacggtc	cggatgtacc	ccttccagat	1200
tcacagcatc	gctctctcca	cctttgcctc	gctcattggc	cccttgggag	gattcttcgc	1260
aagtgaattc	aaacagacct	ttaaaatcaa	agactttgcc	aataccattc	ctggccatgg	1320
aggcatcatg	gatcgctttg	actgccagta	tctgatggcc	acctttgtca	atgtatacat	1380
cgccagtttt	atcagaggcc	ctaaccacaag	caaactgatt	cagcagttcc	tgactttacg	1440
gccagatcac	cagctccaca	tcttcaacac	gctgcggctc	catctgatcg	acaaagggat	1500
gctgacatcc	accacagagg	acgagtaggg	gccaccagg	gccaggagaa	caggaacaga	1560
actgagcagg	ggcagggtctc	caaggcaagc	ccagctgggtg	tgacttagac	aatgacgagg	1620
cttcaactca	ctgtcttttt	tttttttttt	tggagggtat	tttttatttg	tgggttcaaa	1680
aaatctgtat	atacagtccta	tgtgttttaga	atttgtgttg	taagtaaact	acagctttga	1740
gttggaaga	agtcacgggt	tgtaaaacca	tttgattttt	tttaaaacaa	aagtattaat	1800
aatctggaag	acggt					1815

&lt;210&gt; 77

&lt;211&gt; 1740

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3706809CB1

&lt;400&gt; 77

tgacgcagcc	cggtgtctcag	ggaacatggc	ggcgtggtg	agaccgcgca	ggtttgtcgt	60
gcgaccgttg	ctgcagggtg	tccaggcttg	ggaccttgac	gcgaggcgct	gggtccgggc	120
gctgcggcgg	agcccagtg	aagtgggtgt	tccttccgga	gaggtggtgg	aacagaagcg	180
cgctcctggg	aagcagcccc	gcaaggcacc	atctgaggcc	agtgccagg	agcaacgaga	240
gaaacaaccg	ctcgaggagt	ccgcattccc	cgctcccagg	acctgggaag	agtctgggct	300
tcgctacgat	aaagcttatc	ccggggacag	gaggctgagc	agtgtaatga	caatagtaaa	360
gtccaggcca	tttcgggaaa	aacaaggga	gatcctgctg	gaaggtcgca	ggctcatttc	420
agacgctctc	aaggctggag	ctgtgccaaa	aatgttcttc	tttagccgtc	tagaataacct	480
aaaggagtgt	ccagtcgata	agctgaaagg	tgctagcctc	attaagggtga	aatttgagga	540
tatcaaggat	tggtccgacc	tcgtaacgcc	acaaggaata	atggggattt	ttgccaagcc	600
tgaccatggt	aagatgacat	atccaaagac	tcagcttcag	cattcactgc	ctttattatt	660
gatttgtgac	aatctccgtg	accctgggaa	cctggggaca	attctgagat	ctgcagctgg	720
ggcaggctgc	agcaaaagtgt	tactaccaca	aggctgtgtg	gatgectggg	agcccaaagt	780
gctccggggc	ggatggggcg	cacatttccg	gatgccatt	atcaataatc	tgggaatggga	840
aaccgtgccc	aattacctgc	cccctgacac	tcgggtctat	gtggctgaca	actgtggcct	900
ttatgcccag	gctgagatgt	ctaataaagc	tagtgaccat	ggctgggtgt	gtgatcaacg	960
agtgatgaag	tttcacaagt	atgaggaaga	ggaagatgta	gaaaccggag	ccagtcaaga	1020

ttggctgcct	catgttgagg	ttcagagtta	cgactcggac	tggacagagg	cgccggcagc	1080
tgtggtgatt	ggcggggaga	cctacggcgt	gagcctggag	tccttgacag	tggccgagag	1140
cactggtggc	aagaggctgc	tgatccccgt	tgtgcctggt	gtggacagcc	tcaactcggc	1200
catggcggca	agcatcctgc	ttttcgaagg	gaaaagacag	ctgcggggga	ggcgaggaga	1260
cttgagcagg	gacaggagtt	accactgagg	acgcagaagt	gacttctgct	tgaggacgtc	1320
tgcagctcct	cctacaccag	cacactgggtg	ggaggctggc	ggagtcagtg	actatggccc	1380
ccacgttcag	gaggaagggtg	tgatgccgtc	atacagttac	aggaaaaata	agaacttcct	1440
cagaaagaac	aggtcggaat	tcttctgtgc	gcgtcactga	ttttgagggt	ctttttctc	1500
ttggtgacaa	taggtgaccc	acgtggctct	gtgtgttttt	aaaaattgtc	caccaagaag	1560
cactttgtgc	ccagaaagtt	cctgaagcat	catcctggca	gggaggcgcc	tgctccacca	1620
gctggtgggt	gtttgtaatc	gccaaagcacc	agctataggt	cacagccaca	tcactcacag	1680
ctgatcactg	gttggtggaa	aataaactat	gagcagcaga	ttacgttaaa	aaaaaaaaaa	1740

&lt;210&gt; 78

&lt;211&gt; 1225

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3745914CB1

&lt;400&gt; 78

cggggaagagc	cgccgaagc	gtggcggcca	cagactgtgg	gtaccgggtc	cgaggggactc	60
gcgctttttct	ctccgtgcc	tggcgccagc	gaaagccacg	aacgtggtgc	ggctgctact	120
aggctccaca	gcgtgtggc	tttcgcagct	cggctccggg	acggctgccg	cgtccaagtc	180
ggtgactgcc	cacttgccgc	cgaagtggcc	cgagaccccg	ctgctgctgg	aggcaagtga	240
atattatggca	gaagaaagta	atgaaaaatt	ttggcagttt	ttggaaactg	tgcaagaatt	300
agcaattttat	aagcaaacag	aatcagatta	ttcttattac	aacttaatcc	tgaagaaagc	360
tggacagttt	ctagacaatt	tacacatcaa	ccttttaaag	tttgctttct	ctataagggc	420
atactcccca	gctattcaga	tgtttcagca	gattgcagct	gatgagccac	caccagatgg	480
ttgtaatgca	tttgtgtgta	ttcataagaa	gcacacctgt	aaaattaatg	agattaaaaa	540
gctgctgaag	aaagctgctt	caaggactag	accttatcta	tttaaaggag	atcacaaatt	600
tcctacaaac	aaagagaact	taccagtgggt	gattctctat	gccgaaatgg	gtactagaac	660
athtagtgca	tttcacaaag	tattgtctga	aaaagctcaa	aatgaggaaa	ttctgtatgt	720
tcttcgccat	tatattcaga	aaccaagctc	acggaaaatg	tacttatctg	ggtatgggtg	780
ggagctagca	attaagagta	cagaatacaa	agcactggat	gatacccaag	ttaaaactgt	840
gactaatact	actgtagagg	atgagactga	aacaaatgaa	gttcaaggat	ttctctttgg	900
gaaactaaaa	gaaatatatt	cagatcttag	agataatctg	acagcattcc	aaaaatacct	960
gattgagagt	aacaaacaaa	tgatgccttt	gaaagtctgg	gaactacaag	atcttagttt	1020
tcaagcagct	tctcaataaa	tgctcgctcc	agttttatgat	gccattaaat	taatgaaaga	1080
catttcacag	aacttcccca	taaaagccag	agtcctaaatg	attggtaatg	tcttaattgg	1140
atgaatatgt	tgtggagtac	ttttttgcc	agaggatgtc	tcgttgaact	gcttccatga	1200
atactgatgt	tacattaaac	atata				1225

&lt;210&gt; 79

&lt;211&gt; 800

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 4000776CB1

&lt;400&gt; 79

atttacagct	tgatcatcgtg	tcattcccttt	aaccaacatg	tgttttgttg	gtatcaaatg	60
tgtcaggttt	aattgacatt	ttaaaatgtc	ttcaaaaaga	tcacactatg	attcagcatt	120
gaaacgaaaa	gttattgtgt	atgcagaaaa	gcatggaaac	agagcagcag	ggcgtacatt	180
tgaattatgt	gaagcaataa	ttcgtcgtcg	gagaaatgat	cgcaattcca	tattttcctg	240
caaagcaaca	acaaaatggt	ttacggggacc	taagaaagggt	aggtaccac	aagtagatga	300
agctgtacta	cgtttcgtca	gtgagacacg	tgcaaaaagg	ttgcctatca	cacgccaagc	360
aatgcaattg	aaggcaggag	aagttgccaa	aaccttagga	attgatgaaa	caaaattcaa	420
agccacaaga	ggctgggtgtg	accgattcat	gcgtcgagca	ggactatcat	taaggcatca	480
aacatcgttt	tgtcccaagc	tccccactgc	cattaaacag	aagacagtgt	tggagcactc	540
tttcaagaag	tgtgcataa	ccagtactct	tgacaacacc	gggagagacg	ttctgtggaa	600
aaatgcagac	atcaatgact	gtggtttgaa	aagtgtattca	gaagagtggg	attcagaata	660
tgaagttata	attataactt	aaccaattta	tttcacacat	tttctttaca	tatgcacaag	720

attgatgtga taaaaatctg tttaactcaa agcgctgttt caataaatat aaacattttc 780  
tgtgatatga aaaaaaaaaa 800

<210> 80  
<211> 1163  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 4071304CB1

<400> 80  
atcaaaattt actctgcgag actctgctga actggctgca tcagggagaa atttcatctc 60  
ccagaagggt gttgctcatc gttttctccc ggaaacatct gcagagacta gcttttcagg 120  
ctaagggtatc ctccatgatg ttaccattgc aagggtgccca gatgctgcag atgctggaga 180  
aatccttgag gaagagcctc ccagcatect taaaggttta tggaaactgtc ttccacataa 240  
accatgggaaa tccattcaat ctgaaggctg tgggtggacaa gtggcctgat tttaatagag 300  
tggttgctctg ccctcaggag caggatatga cagatgacct tgatcactat accaatactt 360  
accaaattcta ctccaaagat ccccaaaaact gtcaggaatt ccttggatca ccagaactca 420  
tcaactggaa acagcattta cagattcaaa gttcacagcc tagcctgaat gaggctatac 480  
aaaatcttgc agcatttaag tccttcaaa gttcacagcc tagcctgaat gaggctatac 480  
cagctgaaac agccaaggaa ctgactcctt tctctgctgaa atcaaagatt ttatctccca 600  
gtggtggcaa acccaaggcc atcaaccaag agatgtttaa actctcatcc atggatgta 660  
cccattgctca ctggtgtaat aaattctggg attttggtgg taatgagagg agccagagat 720  
tcattgagcg ctgctattcag acctttccca cctgctgtct cctggggcct gaggggaccc 780  
ctgtgtgctg ggatctaag gaccagactg gagagatgag aatggcaggc accttcgagg 840  
aataccgctc ccatggcctt gtgacgtatg tcatctattc ccacgcccag aaattgggca 900  
aacttgggtt tctgtcttat tctcatgtag actacagcaa tgaagctatg caaaaaatga 960  
gttacacact gcaacatgtt cccattccca gaagctggaa ccagtggaa tgtgtacctc 1020  
tgtgatgcca atcctgaaca taagacagtg ttgggcagggt ctgggcgtat aatttgagga 1080  
gtggatggtg gatgggaaga attaattaag cggtcgcaag cttattccct ttagtgaggg 1140  
ttaattttag cttgcaggct acg 1163

<210> 81  
<211> 888  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 4344970CB1

<400> 81  
aaggactatg gcaggggaga acttgcgtac gccgttccac gggcacgtgg gccgcggcgc 60  
cttcagcgac gtgtacgagc ccgcgaggga cacgtttctg cttttggacg cgctcgaggc 120  
agcggtgccc gaactggcag gagtggaaat atgcctggaa gtagggtcag ggtctggtgt 180  
agtatctgca ttcttagcct ctatgatagg ccctcaggct ttgtacatgt gcactgatat 240  
caaccctgag gcagcagctt gtaccctaga gacagcacgc tgaacaaaag ttcacattca 300  
accagttatt acagatttgg tcaaaggctt gctaccaaga ttgaccgaaa aagttgatct 360  
tctggtgttt aatccccctt atgtagtgac tccacctcaa gaggtaggaa gtcacggaat 420  
agaggcagct tgggctggtg gcagaaatgg tcgggaagtc atggacagggt tttttccct 480  
ggttccagat ctccctttcac caagaggatt attctattta gttaccatta aagaaaacaa 540  
cccagaagaa attttgaaaa taatgaagac aaaaggctct caaggaacca ctgcactttc 600  
cagacaagca ggccaagaaa ctctttcagt cctcaagtc accaagtctt agcatacagt 660  
gtgtgcccag aactactgga aactgaatgc atttagcata ttttgaaact gaagtcattc 720  
attaggtaac aaggaatttt atcagaaatt tgtcattaaa aaaaggtaga aagggtggaac 780  
attccctttc agtcaggcct atagaattat ttgcacaagt acaagtaaat atgtgtatta 840  
tatttatagt ttgcatttta aatatagaaa ttatgtatga aaaaaaaa 888

<210> 82  
<211> 1163  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature

&lt;223&gt; Incyte ID No: 5392302CB1

&lt;400&gt; 82

```

gtggcgctctc tgggttacggg gtcggggcaaa gggcagaggt caccgtccag gttggacagc 60
agcacctttg agcgatggcg gcgtctgggg aaccccagag gcagtggcaa gaggagggtg 120
cggcggttgg agtggtgggc tcttgcata cgcacctggt cagtcttact tctcgtttgc 180
caaaaactgg agaaaccatc catggacata agttttttat tggctttgga gggaaagggtg 240
ccaaccagtg tgtccaagct gctcggcttg gagcaatgac gtccatgggtg tgtaagggtg 300
gcaaagattc ttttggcaat gattatatag aaaacttaaa acagaatgat atttctacag 360
aatttacata tcagactaaa gatgctgcta caggaaactgc ttctataatt gtcaataatg 420
aaggccagaa tatcattgtc atagtggctg gagcaaattt acttttgaat acggaggatc 480
tgagggcagc agccaatgtc attagcagag ccaaagtcac ggtctgccag ctcgaaataa 540
ctccagcaac ttctttggaa gccctaacaa tggcccgag gagtggagtg aaaaccttgt 600
tcaatccagc ccttgccatt gctgacctgg atccccagtt ctacaccctc tcagatgtgt 660
tctgctgcaa tgaaagttag gctgagattt taactggcct cacggtgggc agcgctgcag 720
atgctgggga ggctgcatta gtgctcttga aaaggggctg ccagggtggta atcattacct 780
taggggctga aggatgtgtg gtgctgtcac agacagaacc tgagccaaag cacattccca 840
cagagaaagt caaggctgtg gataccacgg gtgctgggtg cagctttgtg ggagctctgg 900
ccttctacct ggcttactat ccaaactctg ccttgggaaga catgctcaac agatccaatt 960
tcattgcagc agtcagtgtc caggctgcag gaacacagtc atcttaccct tacaataaaag 1020
accttcgctg tactctgttt tgattgctat tagtcccaaa ataaatatac ctgggaataa 1080
aatgtacttg ggggtggtg ctcttggtta atgtctatta gaaaatgtcc tctgtccgta 1140
tctttgcaaa tattagtatt ggg                                     1163

```

&lt;210&gt; 83

&lt;211&gt; 931

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5555235CB1

&lt;400&gt; 83

```

gcgaggacag aggtgggctc aacaccaaga ggccaccaat gaaaacgtga ctttttcaca 60
ttccccgccc ctcccggtgt cacgctcccc acgaggggct ggttcctgaa aggatgtaca 120
gaggatcccc aaccgcctgc gaaacccaag ccgcccgtga ggagcgtgcg ttcgggccc 180
cttctcccaac ctgttcgact ccccatcccc aggatgtcaa cctcagtcoc tcaaggccat 240
acctggaccc aacgggtgaa gaaagacgat gagggaggag acccgctgga ccagctgatc 300
tcccgcctctg gctgtgctgc ctcccacttt gcagtgcagg agtgcattgg ccagcaccag 360
gactggcgcc aatgccagcc acaggtgcag cgttcaagg attgcatgag tgaacagcag 420
gcgagggcgg aagaggagct gcagaggagg caagaacaag ccggtgccca ccactgagac 480
cccaaacacc ctatccccag tagatggccc tgccaagacc agcaccagc aagattatag 540
aggaagaaat cctaaatgct ggtgtgggag gtctaaaaa tggggagagt ttttggatct 600
ggagttgaga gccatgggtt tggacatgac tggcacaac agctgtcata tgttcattgt 660
cagatgtcat acattctcag ctgtctgttt ccaccagtat ttaccaggaa aacaaagat 720
gtgttaaggg atgctcccc accccacatc ttaagtcagt gtgccaagta ctgagatgat 780
tttagggaca ttttatttta aattaaattt acaatctaag ggtaaattga ttacctaatt 840
agtgccttct tctttcatca cttgttctgg atgtttcaat cagtttagtg gaaggaaaaa 900
taaattggga aactttttat tcaaaaaaaa a                                     931

```

&lt;210&gt; 84

&lt;211&gt; 1794

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5573296CB1

&lt;400&gt; 84

```

agctcttata aactagtggc aatttctgaa ccagccggc tccatctcag cttctggttt 60
ctaagtccat gtgccaaagg ctgccaggaa ggagacgct tcctgagtc tggatctttc 120
ttccttcttg aaatctttga ctgtgggtag ttatttatt ctgaataaga gcgtccacgc 180
atcatggacc tcgcgggact gctgaagtct cagttcctgt gccacctggg cttctgtctc 240
gtctttattg cctcagggtc aatcatcaac accattcagc tcttactct cctcctctgg 300
cccattaaca agcagctctt ccggaagatc aactgcagac tgtcctattg catctcaagc 360

```



cagctggtga	tgctgctgga	gtggtggtcg	ggcacggaat	gcaccatctt	cacggacccg	420
cgcgccctacc	tcaagtatgg	gaaggaaaat	gccatcgtag	ttctcaacca	caagtttgaa	480
attgactttc	tgtgtggctg	gagcctgtcc	gaacgctttg	ggctgttagg	gggctccaag	540
gtcctggcca	agaaagagct	ggcctatgtc	ccaattatcg	gctggatgtg	gtacttcacc	600
gagatggtct	tctgttcgcg	caagtgggag	caggatcgca	agacggttgc	caccagtttg	660
cagcacctcc	gggactaccc	cgagaagtat	tttttcctga	ttcactgtga	gggcacacgg	720
ttcacggaga	agaagcatga	gatcagcatg	caggtggccc	gggccaaggg	gctgcctcgc	780
ctcaagcatc	acctgttgcc	acgaaccaag	ggcttcgcca	tcaccgtgag	gagcttgaga	840
aatgtagttt	cagctgtata	tgactgtaca	ctcaatttca	gaaataatga	aaatccaaca	900
ctgctgggag	tcctaaacgg	aaagaaatac	catgcagatt	tgtatgttag	gaggatcccc	960
ctggaagaca	tccttgaaga	cgatgacgag	tgctcgccct	ggctgcacaa	gctctaccag	1020
gagaaggatg	cctttcagga	ggagtactac	aggacgggca	ccttcccaga	gacgcccacg	1080
gtgccccccc	ggcgccctcg	gaccctcgtg	aactggctgt	tttgggcctc	gctggtgctc	1140
taccctttct	tccagttcct	ggtcagcatg	atcaggagcg	ggtcttccct	gacgctggcc	1200
agcttcaccc	tcgtcttctt	tgtggcctct	gtgggagttc	gatggatgat	tggtgtgacg	1260
gaaattgaca	agggctctgc	ctacggcaac	tctgacagca	agcagaaact	gaatgactga	1320
ctcagggagg	tgtcaccatc	cgaagggaac	cttggggaac	tggtggcctc	tgcatatcct	1380
ccttagtggg	acacggtgac	aaaggctggg	tgagcccctg	ctgggcacgg	cggaaagtcac	1440
gacctctcca	gccagggagt	ctggtctcaa	ggccggatgg	ggaggaagat	gttttgtaat	1500
ctttttttcc	ccatgtgctt	tagtgggctt	tggttttctt	tttgtgagag	tggtgtgag	1560
aatggctgtg	tggtgagtg	gaactttgtt	ctgtgatcat	agaaagggtg	ttttaggctg	1620
caggggaggg	cagggctggg	gaccgaaggg	gacaagttcc	cctttcatcc	tttgggtgctg	1680
agttttctgt	aacccttgg	tgccagagat	aaagtgaata	gtgcttttagg	tgagatgact	1740
aaattatgcc	tccaagaaaa	aaaaattaaa	gtgcttttct	gggtcaaaaa	aaaa	1794

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
10 May 2001 (10.05.2001)

PCT

(10) International Publication Number  
**WO 01/32888 A3**

(51) International Patent Classification<sup>7</sup>: C12N 15/54,  
9/10, 9/12, A01K 67/00, C07K 16/40, A61K 38/45, G01N  
33/50, C12Q 1/68

(21) International Application Number: PCT/US00/30485

(22) International Filing Date:  
2 November 2000 (02.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/163,595 4 November 1999 (04.11.1999) US

(63) Related by continuation (CON) or continuation-in-part  
(CIP) to earlier application:  
US 60/163,595 (CIP)  
Filed on 4 November 1999 (04.11.1999)

(71) Applicant (for all designated States except US): INCYTE  
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo  
Alto, CA 94030 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom  
[CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).  
YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale,  
CA 94087 (US). HILLMAN, Jennifer, L. [US/US]; 230  
Monroe Drive, #17, Mountain View, CA 94040 (US).  
LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara,  
CA 95054 (US). BANDMAN, Olga [US/US]; 366 Anna  
Avenue, Mountain View, CA 94043 (US). PATTERSON,  
Chandra [US/US]; 490 Sherwood Way, #1, Menlo Park,

CA 94025 (US). SHIH, Leo, L. [US/US]; Apartment  
B, 1081 Tanland Drive, Palo Alto, CA 94303 (US). AZ-  
IMZAI, Yalda [US/US]; 5518 Boulder Canyon Drive,  
Castro Valley, CA 94552 (US). LU, Dyung, Aina, M.  
[US/US]; 233 Coy Drive, San Jose, CA 95123 (US).  
BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road,  
San Leandro, CA 94577 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,  
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:  
— with international search report

(88) Date of publication of the international search report:  
7 March 2002

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: HUMAN TRANSFERASE MOLECULES

(57) Abstract: The invention provides human transferase molecules (HTFS) and polynucleotides which identify and encode HTFS. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HTFS.

WO 01/32888 A3

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/30485

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/54 C12N9/12 A01K67/00 C07K16/40  
A61K38/45 G01N33/50 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N A01K C07K A61K G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EMBL

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>INHAENG YUH ET AL: "Up-regulated uridine kinase gene identified by RLCS in the ventral horn after crush injury to rat sciatic nerves"</p> <p>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS., vol. 266, 1999, pages 104-109, XP000990179</p> <p>ACADEMIC PRESS INC. ORLANDO, FL., US</p> <p>ISSN: 0006-291X</p> <p>the whole document</p> <p style="text-align: center;">--- -/--</p>	1,3, 12-14

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

## \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

26 March 2001

Date of mailing of the international search report

28.06.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

ESPEN, J

## INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 00/30485

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL [Online]  Database entry Q92528, AC Q92528,  1 February 1997 (1997-02-01),  OZAKI K ET AL: "5'-terminal region of UMK"  XP002163961  AA sequence</p>	1
X	<p>-&amp; DATABASE EMBL [Online]  Database entry HSD335, AC D78335,  1 November 1996 (1996-11-01)  OZAKI K ET AL: "Human mRNA for 5'-terminal  region of UMK"  XP002163962  cDNA sequence</p>	1,3, 12-14
X	<p>---  DATABASE EMBL [Online]  Database entry A11741539, AC A1741539,  28 June 1999 (1999-06-28)  NATIONAL CANCER INSTITUTE, CANCER GENOME  ANATOMY PROJECT (CGAP): "Homo sapiens cDNA  clone, 5'-terminal region of UMK"  XP002163963  EST sequence</p>	11,12
Y	<p>---  ROPP P A ET AL: "Cloning and expression  of a cDNA encoding uridine kinase from  mouse brain"  ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS,  vol. 336, 1996, pages 105-112, XP000992365  NEW YORK, US, US  ISSN: 0003-9861  figure 1</p>	1,11-14
Y	<p>---  DATABASE EMBL [Online]  Database entry AW141796; AC AW141796,  2 November 1999 (1999-11-02)  LEE N H ET AL: "Bento Soares Rattus cDNA  clone; 5' end similar to uridine kinase"  XP002163964  EST sequence</p>	1,11-14
Y	<p>---  AHMED N K ET AL: "SOME PROPERTIES OF  URIDINE CYTIDINE KINASE EC-2.7.1.48 FROM A  HUMAN MALIGNANT LYMPHOMA"  CANCER RESEARCH,  vol. 39, no. 8, 1979, pages 3102-3106,  XP000992407  ISSN: 0008-5472  abstract</p> <p>-----</p>	1,11-14

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 00/30485

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
  
Although claim 18 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.: 20,21,23,24  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
  
in part 1-19, 22, 25-28

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1; Claims: in part 1-19,22,25-28; all as far as applicable

Polypeptide being a human transferase molecule and relating to SEQ ID NO 1, polynucleotide coding for said transferase and relating to SEQ ID NO 43; cell transformed with said polynucleotide; transgenic organism comprising said polynucleotide; method for producing said polypeptide; antibody binding to said polypeptide; methods for detecting a target nucleotide in a sample; pharmaceutical composition comprising said polypeptide; methods for screening for an agonist/antagonist of said polypeptide; method for screening for a compound binding to said polypeptide; method for screening for a compound that modulates the activity of said polypeptide; method for screening for a compound that alters the expression of a target nucleotide; method for assessing toxicity of a test compound

Inventions 2-42; Claims: in part 1-19,22,25-28; all as far as applicable

As invention 1 but limited to subject-matter relating to SEQ ID NOs 2-42 and 44-84;

wherein invention 2 is limited to SEQ ID NOs 2 and 44;

wherein invention 3 is limited to SEQ ID NOs 3 and 45; etc.

and invention 42 is limited to SEQ ID NOs 42 and 84.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20,21,23,24

Claims 20 and 23 refer to an agonist/antagonist of the polypeptide of claim 1 without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the result to be achieved. The above comment also applies to claims 21 and 24.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.